



Radomir [YU/US]; 762 Haverhill Drive, Sunnyvale, CA 94068 (US). **DICKSON, Mark** [US/US]; 1411 Gabilan Drive #B, Hollister, CA 95025 (US). **DRMANAC, Snezana** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US). **LABAT, Ivan** [YU/US]; 140 Acalanes Drive, Sunnyvale, CA 94086 (US). **LESHKOWITZ, Dena** [US/US]; 678 Durshire Way, Sunnyvale, CA 94087 (US). **KITA, David** [US/US]; 899 Bounty Drive, Foster City, CA 94404 (US). **GARCIA, Veronica** [ES/US]; 396 Ano Nuevo, Apt. 412, Sunnyvale, CA 94086 (US). **JONES, Lee, William** [US/US]; 396 Ano Nuevo #412, Sunnyvale, CA 94086 (US). **STACHE-CRAIN, Birgit** [DE/US]; 345 South Mary Avenue, Sunnyvale, CA 94086 (US).

(74) **Agent: BLACKBURN, Robert, P.**; Chief Patent Counsel, Chiron Corporation, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,

CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

HUMAN GENES AND GENE EXPRESSION PRODUCTS XVI

Field of the Invention

The present invention relates to polynucleotides of human origin and the encoded gene products.

Background of the Invention

Identification of novel polynucleotides, particularly those that encode an expressed gene product, is important in the advancement of drug discovery, diagnostic technologies, and the understanding of the progression and nature of complex diseases such as cancer. Identification of genes expressed in different cell types isolated from sources that differ in disease state or stage, developmental stage, exposure to various environmental factors, the tissue of origin, the species from which the tissue was isolated, and the like is key to identifying the genetic factors that are responsible for the phenotypes associated with these various differences.

This invention provides novel human polynucleotides, the polypeptides encoded by these polynucleotides, and the genes and proteins corresponding to these novel polynucleotides.

Summary of the Invention

This invention relates to novel human polynucleotides and variants thereof, their encoded polypeptides and variants thereof, to genes corresponding to these polynucleotides and to proteins expressed by the genes. The invention also relates to diagnostics and therapeutics comprising such novel human polynucleotides, their corresponding genes or gene products, including probes, antisense nucleotides, and antibodies. The polynucleotides of the invention correspond to a polynucleotide comprising the sequence information of at least one of SEQ ID NOS:1-316.

Various aspects and embodiments of the invention will be readily apparent to the ordinarily skilled artisan upon reading the description provided herein.

Brief Description of the Figures

Figures 1A-1B is a comparison of SEQ ID NO:315 and clone H72034 (SEQ ID NO:317).

Figure 2 is a comparison of SEQ ID NO:316 and clone AA707002 (SEQ ID NO:318).

Detailed Description of the Invention

The invention relates to polynucleotides comprising the disclosed nucleotide sequences, to full length cDNA, mRNA genomic sequences, and genes corresponding to these sequences and degenerate variants thereof, and to polypeptides encoded by the polynucleotides of the invention and polypeptide variants. The following detailed description describes the polynucleotide compositions encompassed by the invention, methods for obtaining cDNA or genomic DNA

encoding a full-length gene product, expression of these polynucleotides and genes, identification of structural motifs of the polynucleotides and genes, identification of the function of a gene product encoded by a gene corresponding to a polynucleotide of the invention, use of the provided polynucleotides as probes and in mapping and in tissue profiling, use of the corresponding
5 polypeptides and other gene products to raise antibodies, and use of the polynucleotides and their encoded gene products for therapeutic and diagnostic purposes.

Polynucleotide Compositions

The scope of the invention with respect to polynucleotide compositions includes, but is not
10 necessarily limited to, polynucleotides having a sequence set forth in any one of SEQ ID NOS:1-316; polynucleotides obtained from the biological materials described herein or other biological sources (particularly human sources) by hybridization under stringent conditions (particularly conditions of high stringency); genes corresponding to the provided polynucleotides; variants of the provided polynucleotides and their corresponding genes, particularly those variants that retain a
15 biological activity of the encoded gene product (*e.g.*, a biological activity ascribed to a gene product corresponding to the provided polynucleotides as a result of the assignment of the gene product to a protein family(ies) and/or identification of a functional domain present in the gene product). Other nucleic acid compositions contemplated by and within the scope of the present invention will be readily apparent to one of ordinary skill in the art when provided with the disclosure here.
20 “Polynucleotide” and “nucleic acid” as used herein with reference to nucleic acids of the composition is not intended to be limiting as to the length or structure of the nucleic acid unless specifically indicated.

The invention features polynucleotides that are expressed in human tissue, specifically human colon, breast, and/or lung tissue. Novel nucleic acid compositions of the invention of
25 particular interest comprise a sequence set forth in any one of SEQ ID NOS:1-316 or an identifying sequence thereof. An “identifying sequence” is a contiguous sequence of residues at least about 10 nt to about 20 nt in length, usually at least about 50 nt to about 100 nt in length, that uniquely identifies a polynucleotide sequence, *e.g.*, exhibits less than 90%, usually less than about 80% to about 85% sequence identity to any contiguous nucleotide sequence of more than about
30 20 nt. Thus, the subject novel nucleic acid compositions include full length cDNAs or mRNAs that encompass an identifying sequence of contiguous nucleotides from any one of SEQ ID NOS: 1-316.

The polynucleotides of the invention also include polynucleotides having sequence similarity or sequence identity. Nucleic acids having sequence similarity are detected by
35 hybridization under low stringency conditions, for example, at 50°C and 10XSSC (0.9 M saline/0.09 M sodium citrate) and remain bound when subjected to washing at 55°C in 1XSSC.

Sequence identity can be determined by hybridization under stringent conditions, for example, at 50°C or higher and 0.1XSSC (9 mM saline/0.9 mM sodium citrate). Hybridization methods and conditions are well known in the art, see, *e.g.*, USPN 5,707,829. Nucleic acids that are substantially identical to the provided polynucleotide sequences, *e.g.* allelic variants, genetically altered versions of the gene, *etc.*, bind to the provided polynucleotide sequences (SEQ ID NOS:1-316) under stringent hybridization conditions. By using probes, particularly labeled probes of DNA sequences, one can isolate homologous or related genes. The source of homologous genes can be any species, *e.g.* primate species, particularly human; rodents, such as rats and mice; canines, felines, bovines, ovines, equines, yeast, nematodes, *etc.*

Preferably, hybridization is performed using at least 15 contiguous nucleotides (nt) of at least one of SEQ ID NOS:1-316. That is, when at least 15 contiguous nt of one of the disclosed SEQ ID NOS. is used as a probe, the probe will preferentially hybridize with a nucleic acid comprising the complementary sequence, allowing the identification and retrieval of the nucleic acids that uniquely hybridize to the selected probe. Probes from more than one SEQ ID NO. can hybridize with the same nucleic acid if the cDNA from which they were derived corresponds to one mRNA. Probes of more than 15 nt can be used, *e.g.*, probes of from about 18 nt to about 100 nt, but 15 nt represents sufficient sequence for unique identification.

The polynucleotides of the invention also include naturally occurring variants of the nucleotide sequences (*e.g.*, degenerate variants, allelic variants, *etc.*). Variants of the polynucleotides of the invention are identified by hybridization of putative variants with nucleotide sequences disclosed herein, preferably by hybridization under stringent conditions. For example, by using appropriate wash conditions, variants of the polynucleotides of the invention can be identified where the allelic variant exhibits at most about 25-30% base pair (bp) mismatches relative to the selected polynucleotide probe. In general, allelic variants contain 15-25% bp mismatches, and can contain as little as even 5-15%, or 2-5%, or 1-2% bp mismatches, as well as a single bp mismatch.

The invention also encompasses homologs corresponding to the polynucleotides of SEQ ID NOS:1-316, where the source of homologous genes can be any mammalian species, *e.g.*, primate species, particularly human; rodents, such as rats; canines, felines, bovines, ovines, equines, yeast, nematodes, *etc.* Between mammalian species, *e.g.*, human and mouse, homologs generally have substantial sequence similarity, *e.g.*, at least 75% sequence identity, usually at least 90%, more usually at least 95% between nucleotide sequences. Sequence similarity is calculated based on a reference sequence, which may be a subset of a larger sequence, such as a conserved motif, coding region, flanking region, *etc.* A reference sequence will usually be at least about 18 contiguous nt long, more usually at least about 30 nt long, and may extend to the complete

sequence that is being compared. Algorithms for sequence analysis are known in the art, such as gapped BLAST, described in Altschul, et al. *Nucleic Acids Res.* (1997) 25:3389-3402.

In general, variants of the invention have a sequence identity greater than at least about 65%, preferably at least about 75%, more preferably at least about 85%, and can be greater than at least about 90% or more as determined by the Smith-Waterman homology search algorithm as implemented in MPSRCH program (Oxford Molecular). For the purposes of this invention, a preferred method of calculating percent identity is the Smith-Waterman algorithm, using the following. Global DNA sequence identity must be greater than 65% as determined by the Smith-Waterman homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with the following search parameters: gap open penalty, 12; and gap extension penalty, 1.

The subject nucleic acids can be cDNAs or genomic DNAs, as well as fragments thereof, particularly fragments that encode a biologically active gene product and/or are useful in the methods disclosed herein (*e.g.*, in diagnosis, as a unique identifier of a differentially expressed gene of interest, *etc.*). The term "cDNA" as used herein is intended to include all nucleic acids that share the arrangement of sequence elements found in native mature mRNA species, where sequence elements are exons and 3' and 5' non-coding regions. Normally mRNA species have contiguous exons, with the intervening introns, when present, being removed by nuclear RNA splicing, to create a continuous open reading frame encoding a polypeptide of the invention.

A genomic sequence of interest comprises the nucleic acid present between the initiation codon and the stop codon, as defined in the listed sequences, including all of the introns that are normally present in a native chromosome. It can further include the 3' and 5' untranslated regions found in the mature mRNA. It can further include specific transcriptional and translational regulatory sequences, such as promoters, enhancers, *etc.*, including about 1 kb, but possibly more, of flanking genomic DNA at either the 5' and 3' end of the transcribed region. The genomic DNA can be isolated as a fragment of 100 kbp or smaller; and substantially free of flanking chromosomal sequence. The genomic DNA flanking the coding region, either 3' and 5', or internal regulatory sequences as sometimes found in introns, contains sequences required for proper tissue, stage-specific, or disease-state specific expression.

The nucleic acid compositions of the subject invention can encode all or a part of the subject polypeptides. Double or single stranded fragments can be obtained from the DNA sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, by restriction enzyme digestion, by PCR amplification, *etc.* Isolated polynucleotides and polynucleotide fragments of the invention comprise at least about 10, about 15, about 20, about 35, about 50, about 100, about 150 to about 200, about 250 to about 300, or about 350 contiguous nt selected from the polynucleotide sequences as shown in SEQ ID NOS:1-316. For the most part,

fragments will be of at least 15 nt, usually at least 18 nt or 25 nt, and up to at least about 50 contiguous nt in length or more. In a preferred embodiment, the polynucleotide molecules comprise a contiguous sequence of at least 12 nt selected from the group consisting of the polynucleotides shown in SEQ ID NOS:1-316.

5 Probes specific to the polynucleotides of the invention can be generated using the polynucleotide sequences disclosed in SEQ ID NOS:1-316. The probes are preferably at least about a 12, 15, 16, 18, 20, 22, 24, or 25 nt fragment of a corresponding contiguous sequence of SEQ ID NOS:1-316, and can be less than 2, 1, 0.5, 0.1, or 0.05 kb in length. The probes can be synthesized chemically or can be generated from longer polynucleotides using restriction enzymes.

10 The probes can be labeled, for example, with a radioactive, biotinylated, or fluorescent tag. Preferably, probes are designed based upon an identifying sequence of a polynucleotide of one of SEQ ID NOS:1-316. More preferably, probes are designed based on a contiguous sequence of one of the subject polynucleotides that remain unmasked following application of a masking program for masking low complexity (*e.g.*, XBLAST) to the sequence., *i.e.*, one would select an unmasked

15 region, as indicated by the polynucleotides outside the poly-n stretches of the masked sequence produced by the masking program.

The polynucleotides of the subject invention are isolated and obtained in substantial purity, generally as other than an intact chromosome. Usually, the polynucleotides, either as DNA or RNA, will be obtained substantially free of other naturally-occurring nucleic acid sequences,

20 generally being at least about 50%, usually at least about 90% pure and are typically “recombinant”, *e.g.*, flanked by one or more nucleotides with which it is not normally associated on a naturally occurring chromosome.

The polynucleotides of the invention can be provided as a linear molecule or within a circular molecule, and can be provided within autonomously replicating molecules (vectors) or

25 within molecules without replication sequences. Expression of the polynucleotides can be regulated by their own or by other regulatory sequences known in the art. The polynucleotides of the invention can be introduced into suitable host cells using a variety of techniques available in the art, such as transferrin polycation-mediated DNA transfer, transfection with naked or encapsulated nucleic acids, liposome-mediated DNA transfer, intracellular transportation of DNA-coated latex

30 beads, protoplast fusion, viral infection, electroporation, gene gun, calcium phosphate-mediated transfection, and the like.

The subject nucleic acid compositions can be used to, for example, produce polypeptides, as probes for the detection of mRNA of the invention in biological samples (*e.g.*, extracts of human cells) to generate additional copies of the polynucleotides, to generate ribozymes or antisense

35 oligonucleotides, and as single stranded DNA probes or as triple-strand forming oligonucleotides. The probes described herein can be used to, for example, determine the presence or absence of the

polynucleotide sequences as shown in SEQ ID NOS:1-316 or variants thereof in a sample. These and other uses are described in more detail below.

Use of Polynucleotides to Obtain Full-Length cDNA, Gene, and Promoter Region

Full-length cDNA molecules comprising the disclosed polynucleotides are obtained as follows. A polynucleotide having a sequence of one of SEQ ID NOS:1-316, or a portion thereof comprising at least 12, 15, 18, or 20 nt, is used as a hybridization probe to detect hybridizing members of a cDNA library using probe design methods, cloning methods, and clone selection techniques such as those described in USPN 5,654,173. Libraries of cDNA are made from selected tissues, such as normal or tumor tissue, or from tissues of a mammal treated with, for example, a pharmaceutical agent. Preferably, the tissue is the same as the tissue from which the polynucleotides of the invention were isolated, as both the polynucleotides described herein and the cDNA represent expressed genes. Most preferably, the cDNA library is made from the biological material described herein in the Examples. The choice of cell type for library construction can be made after the identity of the protein encoded by the gene corresponding to the polynucleotide of the invention is known. This will indicate which tissue and cell types are likely to express the related gene, and thus represent a suitable source for the mRNA for generating the cDNA. Where the provided polynucleotides are isolated from cDNA libraries, the libraries are prepared from mRNA of human colon cells, more preferably, human colon cancer cells, even more preferably, from a highly metastatic colon cell, Km12L4-A.

Techniques for producing and probing nucleic acid sequence libraries are described, for example, in Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, (1989) Cold Spring Harbor Press, Cold Spring Harbor, NY. The cDNA can be prepared by using primers based on sequence from SEQ ID NOS:1-316. In one embodiment, the cDNA library can be made from only poly-adenylated mRNA. Thus, poly-T primers can be used to prepare cDNA from the mRNA.

Members of the library that are larger than the provided polynucleotides, and preferably that encompass the complete coding sequence of the native message, are obtained. In order to confirm that the entire cDNA has been obtained, RNA protection experiments are performed as follows. Hybridization of a full-length cDNA to an mRNA will protect the RNA from RNase degradation. If the cDNA is not full length, then the portions of the mRNA that are not hybridized will be subject to RNase degradation. This is assayed, as is known in the art, by changes in electrophoretic mobility on polyacrylamide gels, or by detection of released monoribonucleotides. Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, (1989) Cold Spring Harbor Press, Cold Spring Harbor, NY. In order to obtain additional sequences 5' to the end of a partial cDNA, 5' RACE (*PCR Protocols: A Guide to Methods and Applications*, (1990) Academic Press, Inc.) can be performed.

Genomic DNA is isolated using the provided polynucleotides in a manner similar to the isolation of full-length cDNAs. Briefly, the provided polynucleotides, or portions thereof, are used as probes to libraries of genomic DNA. Preferably, the library is obtained from the cell type that was used to generate the polynucleotides of the invention, but this is not essential. Most preferably, the genomic DNA is obtained from the biological material described herein in the Examples. Such libraries can be in vectors suitable for carrying large segments of a genome, such as P1 or YAC, as described in detail in Sambrook *et al.*, 9.4-9.30. In addition, genomic sequences can be isolated from human BAC libraries, which are commercially available from Research Genetics, Inc., Huntsville, Alabama, USA, for example. In order to obtain additional 5' or 3' sequences, chromosome walking is performed, as described in Sambrook *et al.*, such that adjacent and overlapping fragments of genomic DNA are isolated. These are mapped and pieced together, as is known in the art, using restriction digestion enzymes and DNA ligase.

Using the polynucleotide sequences of the invention, corresponding full-length genes can be isolated using both classical and PCR methods to construct and probe cDNA libraries. Using either method, Northern blots, preferably, are performed on a number of cell types to determine which cell lines express the gene of interest at the highest level. Classical methods of constructing cDNA libraries are taught in Sambrook *et al.*, *supra*. With these methods, cDNA can be produced from mRNA and inserted into viral or expression vectors. Typically, libraries of mRNA comprising poly(A) tails can be produced with poly(T) primers. Similarly, cDNA libraries can be produced using the instant sequences as primers.

PCR methods are used to amplify the members of a cDNA library that comprise the desired insert. In this case, the desired insert will contain sequence from the full length cDNA that corresponds to the instant polynucleotides. Such PCR methods include gene trapping and RACE methods. Gene trapping entails inserting a member of a cDNA library into a vector. The vector then is denatured to produce single stranded molecules. Next, a substrate-bound probe, such as a biotinylated oligo, is used to trap cDNA inserts of interest. Biotinylated probes can be linked to an avidin-bound solid substrate. PCR methods can be used to amplify the trapped cDNA. To trap sequences corresponding to the full length genes, the labeled probe sequence is based on the polynucleotide sequences of the invention. Random primers or primers specific to the library vector can be used to amplify the trapped cDNA. Such gene trapping techniques are described in Gruber *et al.*, WO 95/04745 and Gruber *et al.*, USPN 5,500,356. Kits are commercially available to perform gene trapping experiments from, for example, Life Technologies, Gaithersburg, Maryland, USA.

"Rapid amplification of cDNA ends," or RACE, is a PCR method of amplifying cDNAs from a number of different RNAs. The cDNAs are ligated to an oligonucleotide linker, and amplified by PCR using two primers. One primer is based on sequence from the instant

polynucleotides, for which full length sequence is desired, and a second primer comprises sequence that hybridizes to the oligonucleotide linker to amplify the cDNA. A description of this methods is reported in WO 97/19110. In preferred embodiments of RACE, a common primer is designed to anneal to an arbitrary adaptor sequence ligated to cDNA ends (Apte and Siebert, *Biotechniques* 5 (1993) 15:890-893; Edwards *et al.*, *Nuc. Acids Res.* (1991) 19:5227-5232). When a single gene-specific RACE primer is paired with the common primer, preferential amplification of sequences between the single gene specific primer and the common primer occurs. Commercial cDNA pools modified for use in RACE are available.

Another PCR-based method generates full-length cDNA library with anchored ends 10 without needing specific knowledge of the cDNA sequence. The method uses lock-docking primers (I-VI), where one primer, poly TV (I-III) locks over the polyA tail of eukaryotic mRNA producing first strand synthesis and a second primer, polyGH (IV-VI) locks onto the polyC tail added by terminal deoxynucleotidyl transferase (TdT)(see, e.g., WO 96/40998).

The promoter region of a gene generally is located 5' to the initiation site for RNA 15 polymerase II. Hundreds of promoter regions contain the "TATA" box, a sequence such as TATTA or TATAA, which is sensitive to mutations. The promoter region can be obtained by performing 5' RACE using a primer from the coding region of the gene. Alternatively, the cDNA can be used as a probe for the genomic sequence, and the region 5' to the coding region is identified by "walking up." If the gene is highly expressed or differentially expressed, the promoter from the 20 gene can be of use in a regulatory construct for a heterologous gene.

Once the full-length cDNA or gene is obtained, DNA encoding variants can be prepared by site-directed mutagenesis, described in detail in Sambrook *et al.*, 15.3-15.63. The choice of codon or nucleotide to be replaced can be based on disclosure herein on optional changes in amino acids to achieve altered protein structure and/or function.

As an alternative method to obtaining DNA or RNA from a biological material, nucleic 25 acid comprising nucleotides having the sequence of one or more polynucleotides of the invention can be synthesized. Thus, the invention encompasses nucleic acid molecules ranging in length from 15 nt (corresponding to at least 15 contiguous nt of one of SEQ ID NOS:1-316) up to a maximum length suitable for one or more biological manipulations, including replication and 30 expression, of the nucleic acid molecule. The invention includes but is not limited to (a) nucleic acid having the size of a full gene, and comprising at least one of SEQ ID NOS:1-316; (b) the nucleic acid of (a) also comprising at least one additional gene, operably linked to permit expression of a fusion protein; (c) an expression vector comprising (a) or (b); (d) a plasmid comprising (a) or (b) ; and (e) a recombinant viral particle comprising (a) or (b). Once provided 35 with the polynucleotides disclosed herein, construction or preparation of (a) - (e) are well within the skill in the art.

The sequence of a nucleic acid comprising at least 15 contiguous nt of at least any one of SEQ ID NOS:1-316, preferably the entire sequence of at least any one of SEQ ID NOS:1-316, is not limited and can be any sequence of A, T, G, and/or C (for DNA) and A, U, G, and/or C (for RNA) or modified bases thereof, including inosine and pseudouridine. The choice of sequence will depend on the desired function and can be dictated by coding regions desired, the intron-like regions desired, and the regulatory regions desired. Where the entire sequence of any one of SEQ ID NOS:1-316 is within the nucleic acid, the nucleic acid obtained is referred to herein as a polynucleotide comprising the sequence of any one of SEQ ID NOS:1-316.

Expression of Polypeptide Encoded by Full-Length cDNA or Full-Length Gene

The provided polynucleotides (e.g., a polynucleotide having a sequence of one of SEQ ID NOS:1-316), the corresponding cDNA, or the full-length gene is used to express a partial or complete gene product. Constructs of polynucleotides having sequences of SEQ ID NOS:1-316 can also be generated synthetically. Alternatively, single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides is described by, e.g., Stemmer *et al.*, *Gene (Amsterdam)* (1995) 164(1):49-53. In this method, assembly PCR (the synthesis of long DNA sequences from large numbers of oligodeoxyribonucleotides (oligos)) is described. The method is derived from DNA shuffling (Stemmer, *Nature* (1994) 370:389-391), and does not rely on DNA ligase, but instead relies on DNA polymerase to build increasingly longer DNA fragments during the assembly process.

Appropriate polynucleotide constructs are purified using standard recombinant DNA techniques as described in, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, (1989) Cold Spring Harbor Press, Cold Spring Harbor, NY, and under current regulations described in United States Dept. of HHS, National Institute of Health (NIH) Guidelines for Recombinant DNA Research. The gene product encoded by a polynucleotide of the invention is expressed in any expression system, including, for example, bacterial, yeast, insect, amphibian and mammalian systems. Vectors, host cells and methods for obtaining expression in same are well known in the art. Suitable vectors and host cells are described in USPN 5,654,173.

Polynucleotide molecules comprising a polynucleotide sequence provided herein are generally propagated by placing the molecule in a vector. Viral and non-viral vectors are used, including plasmids. The choice of plasmid will depend on the type of cell in which propagation is desired and the purpose of propagation. Certain vectors are useful for amplifying and making large amounts of the desired DNA sequence. Other vectors are suitable for expression in cells in culture. Still other vectors are suitable for transfer and expression in cells in a whole animal or person. The choice of appropriate vector is well within the skill of the art. Many such vectors are available commercially. Methods for preparation of vectors comprising a desired sequence are well known in the art.

The polynucleotides set forth in SEQ ID NOS:1-316 or their corresponding full-length polynucleotides are linked to regulatory sequences as appropriate to obtain the desired expression properties. These can include promoters (attached either at the 5' end of the sense strand or at the 3' end of the antisense strand), enhancers, terminators, operators, repressors, and inducers. The promoters can be regulated or constitutive. In some situations it may be desirable to use conditionally active promoters, such as tissue-specific or developmental stage-specific promoters. These are linked to the desired nucleotide sequence using the techniques described above for linkage to vectors. Any techniques known in the art can be used.

When any of the above host cells, or other appropriate host cells or organisms, are used to replicate and/or express the polynucleotides or nucleic acids of the invention, the resulting replicated nucleic acid, RNA, expressed protein or polypeptide, is within the scope of the invention as a product of the host cell or organism. The product is recovered by any appropriate means known in the art.

Once the gene corresponding to a selected polynucleotide is identified, its expression can be regulated in the cell to which the gene is native. For example, an endogenous gene of a cell can be regulated by an exogenous regulatory sequence as disclosed in USPN 5,641,670.

Identification of Functional and Structural Motifs of Novel Genes Screening Against Publicly Available Databases

Translations of the nucleotide sequence of the provided polynucleotides, cDNAs or full genes can be aligned with individual known sequences. Similarity with individual sequences can be used to determine the activity of the polypeptides encoded by the polynucleotides of the invention. Also, sequences exhibiting similarity with more than one individual sequence can exhibit activities that are characteristic of either or both individual sequences.

The full length sequences and fragments of the polynucleotide sequences of the nearest neighbors can be used as probes and primers to identify and isolate the full length sequence corresponding to provided polynucleotides. The nearest neighbors can indicate a tissue or cell type to be used to construct a library for the full-length sequences corresponding to the provided polynucleotides.

Typically, a selected polynucleotide is translated in all six frames to determine the best alignment with the individual sequences. The sequences disclosed herein in the Sequence Listing are in a 5' to 3' orientation and translation in three frames can be sufficient (with a few specific exceptions as described in the Examples). These amino acid sequences are referred to, generally, as query sequences, which will be aligned with the individual sequences. Databases with individual sequences are described in "Computer Methods for Macromolecular Sequence Analysis" *Methods in Enzymology* (1996) 266, Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co.,

San Diego, California, USA. Databases include GenBank, EMBL, and DNA Database of Japan (DDBJ).

Query and individual sequences can be aligned using the methods and computer programs described above, and include BLAST 2.0, available over the world wide web at a site supported by the National Center for Biotechnology Information, which is supported by the National Library of Medicine and the National Institutes of Health. See also Altschul, et al. *Nucleic Acids Res.* (1997) 25:3389-3402. Another alignment algorithm is Fasta, available in the Genetics Computing Group (GCG) package, Madison, Wisconsin, USA, a wholly owned subsidiary of Oxford Molecular Group, Inc. Other techniques for alignment are described in Doolittle, *supra*. Preferably, an alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See *Meth. Mol. Biol.* (1997) 70: 173-187. Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAR computer. MPSRCH uses a Smith-Waterman algorithm to score sequences on a massively parallel computer. This approach improves ability to identify sequences that are distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Amino acid sequences encoded by the provided polynucleotides can be used to search both protein and DNA databases. Incorporated herein by reference are all sequences that have been made public as of the filing date of this application by any of the DNA or protein sequence databases, including the patent databases (*e.g.*, GeneSeq). Also incorporated by reference are those sequences that have been submitted to these databases as of the filing date of the present application but not made public until after the filing date of the present application.

Results of individual and query sequence alignments can be divided into three categories: high similarity, weak similarity, and no similarity. Individual alignment results ranging from high similarity to weak similarity provide a basis for determining polypeptide activity and/or structure. Parameters for categorizing individual results include: percentage of the alignment region length where the strongest alignment is found, percent sequence identity, and p value. The percentage of the alignment region length is calculated by counting the number of residues of the individual sequence found in the region of strongest alignment, *e.g.*, contiguous region of the individual sequence that contains the greatest number of residues that are identical to the residues of the corresponding region of the aligned query sequence. This number is divided by the total residue length of the query sequence to calculate a percentage. For example, a query sequence of 20 amino acid residues might be aligned with a 20 amino acid region of an individual sequence. The individual sequence might be identical to amino acid residues 5, 9-15, and 17-19 of the query sequence. The region of strongest alignment is thus the region stretching from residue 9-19, an 11

amino acid stretch. The percentage of the alignment region length is: 11 (length of the region of strongest alignment) divided by (query sequence length) 20 or 55%.

Percent sequence identity is calculated by counting the number of amino acid matches between the query and individual sequence and dividing total number of matches by the number of residues of the individual sequences found in the region of strongest alignment. Thus, the percent identity in the example above would be 10 matches divided by 11 amino acids, or approximately, 90.9%

P value is the probability that the alignment was produced by chance. For a single alignment, the p value can be calculated according to Karlin *et al.*, *Proc. Natl. Acad. Sci.* (1990) 87:2264 and Karlin *et al.*, *Proc. Natl. Acad. Sci.* (1993) 90. The p value of multiple alignments using the same query sequence can be calculated using an heuristic approach described in Altschul *et al.*, *Nat. Genet.* (1994) 6:119. Alignment programs such as BLAST program can calculate the p value. See also Altschul *et al.*, *Nucleic Acids Res.* (1997) 25:3389-3402.

Another factor to consider for determining identity or similarity is the location of the similarity or identity. Strong local alignment can indicate similarity even if the length of alignment is short. Sequence identity scattered throughout the length of the query sequence also can indicate a similarity between the query and profile sequences. The boundaries of the region where the sequences align can be determined according to Doolittle, *supra*; BLAST 2.0 (see, *e.g.*, Altschul, *et al. Nucleic Acids Res.* (1997) 25:3389-3402) or FAST programs; or by determining the area where sequence identity is highest.

High Similarity. In general, in alignment results considered to be of high similarity, the percent of the alignment region length is typically at least about 55% of total length query sequence; more typically, at least about 58%; even more typically; at least about 60% of the total residue length of the query sequence. Usually, percent length of the alignment region can be as much as about 62%; more usually, as much as about 64%; even more usually, as much as about 66%. Further, for high similarity, the region of alignment, typically, exhibits at least about 75% of sequence identity; more typically, at least about 78%; even more typically; at least about 80% sequence identity. Usually, percent sequence identity can be as much as about 82%; more usually, as much as about 84%; even more usually, as much as about 86%.

The p value is used in conjunction with these methods. If high similarity is found, the query sequence is considered to have high similarity with a profile sequence when the p value is less than or equal to about 10^{-2} ; more usually; less than or equal to about 10^{-3} ; even more usually; less than or equal to about 10^{-4} . More typically, the p value is no more than about 10^{-5} ; more typically; no more than or equal to about 10^{-10} ; even more typically; no more than or equal to about 10^{-15} for the query sequence to be considered high similarity.

Weak Similarity. In general, where alignment results considered to be of weak similarity, there is no minimum percent length of the alignment region nor minimum length of alignment. A better showing of weak similarity is considered when the region of alignment is, typically, at least about 15 amino acid residues in length; more typically, at least about 20; even more typically; at least about 25 amino acid residues in length. Usually, length of the alignment region can be as much as about 30 amino acid residues; more usually, as much as about 40; even more usually, as much as about 60 amino acid residues. Further, for weak similarity, the region of alignment, typically, exhibits at least about 35% of sequence identity; more typically, at least about 40%; even more typically; at least about 45% sequence identity. Usually, percent sequence identity can be as much as about 50%; more usually, as much as about 55%; even more usually, as much as about 60%.

If low similarity is found, the query sequence is considered to have weak similarity with a profile sequence when the p value is usually less than or equal to about 10^{-2} ; more usually; less than or equal to about 10^{-3} ; even more usually; less than or equal to about 10^{-4} . More typically, the p value is no more than about 10^{-5} ; more usually; no more than or equal to about 10^{-10} ; even more usually; no more than or equal to about 10^{-15} for the query sequence to be considered weak similarity.

Similarity Determined by Sequence Identity Alone. Sequence identity alone can be used to determine similarity of a query sequence to an individual sequence and can indicate the activity of the sequence. Such an alignment, preferably, permits gaps to align sequences. Typically, the query sequence is related to the profile sequence if the sequence identity over the entire query sequence is at least about 15%; more typically, at least about 20%; even more typically, at least about 25%; even more typically, at least about 50%. Sequence identity alone as a measure of similarity is most useful when the query sequence is usually, at least 80 residues in length; more usually, 90 residues; even more usually, at least 95 amino acid residues in length. More typically, similarity can be concluded based on sequence identity alone when the query sequence is preferably 100 residues in length; more preferably, 120 residues in length; even more preferably, 150 amino acid residues in length.

Alignments with Profile and Multiple Aligned Sequences. Translations of the provided polynucleotides can be aligned with amino acid profiles that define either protein families or common motifs. Also, translations of the provided polynucleotides can be aligned to multiple sequence alignments (MSA) comprising the polypeptide sequences of members of protein families or motifs. Similarity or identity with profile sequences or MSAs can be used to determine the activity of the gene products (*e.g.*, polypeptides) encoded by the provided polynucleotides or corresponding cDNA or genes. For example, sequences that show an identity or similarity with a chemokine profile or MSA can exhibit chemokine activities.

Profiles can be designed manually by (1) creating an MSA, which is an alignment of the amino acid sequence of members that belong to the family and (2) constructing a statistical representation of the alignment. Such methods are described, for example, in Birney *et al.*, *Nucl. Acid Res.* (1996) 24(14): 2730-2739. MSAs of some protein families and motifs are publicly available. For example, the Genome Sequencing Center at the Washington University School of Medicine provides a web set (Pfam) which includes MSAs of 547 different families and motifs. These MSAs are described also in Sonnhammer *et al.*, *Proteins* (1997) 28: 405-420. Other sources over the world wide web include the site supported by the European Molecular Biology Laboratories in Heidelberg, Germany. A brief description of these MSAs is reported in Pascarella *et al.*, *Prot. Eng.* (1996) 9(3):249-251. Techniques for building profiles from MSAs are described in Sonnhammer *et al.*, *supra*; Birney *et al.*, *supra*; and "Computer Methods for Macromolecular Sequence Analysis," *Methods in Enzymology* (1996) 266, Doolittle, Academic Press, Inc., San Diego, California, USA.

Similarity between a query sequence and a protein family or motif can be determined by (a) comparing the query sequence against the profile and/or (b) aligning the query sequence with the members of the family or motif. Typically, a program such as Searchwise is used to compare the query sequence to the statistical representation of the multiple alignment, also known as a profile (see Birney *et al.*, *supra*). Other techniques to compare the sequence and profile are described in Sonnhammer *et al.*, *supra* and Doolittle, *supra*.

Next, methods described by Feng *et al.*, *J. Mol. Evol.* (1987) 25:351 and Higgins *et al.*, *CABIOS* (1989) 5:151 can be used to align the query sequence with the members of a family or motif, also known as a MSA. Sequence alignments can be generated using any of a variety of software tools. Examples include PileUp, which creates a multiple sequence alignment, and is described in Feng *et al.*, *J. Mol. Evol.* (1987) 25:351. Another method, GAP, uses the alignment method of Needleman *et al.*, *J. Mol. Biol.* (1970) 48:443. GAP is best suited for global alignment of sequences. A third method, BestFit, functions by inserting gaps to maximize the number of matches using the local homology algorithm of Smith *et al.*, *Adv. Appl. Math.* (1981) 2:482. In general, the following factors are used to determine if a similarity between a query sequence and a profile or MSA exists: (1) number of conserved residues found in the query sequence, (2) percentage of conserved residues found in the query sequence, (3) number of frameshifts, and (4) spacing between conserved residues.

Some alignment programs that both translate and align sequences can make any number of frameshifts when translating the nucleotide sequence to produce the best alignment. The fewer frameshifts needed to produce an alignment, the stronger the similarity or identity between the query and profile or MSAs. For example, a weak similarity resulting from no frameshifts can be a better indication of activity or structure of a query sequence, than a strong similarity resulting from

two frameshifts. Preferably, three or fewer frameshifts are found in an alignment; more preferably two or fewer frameshifts; even more preferably, one or fewer frameshifts; even more preferably, no frameshifts are found in an alignment of query and profile or MSAs.

Conserved residues are those amino acids found at a particular position in all or some of the family or motif members. Alternatively, a position is considered conserved if only a certain class of amino acids is found in a particular position in all or some of the family members. For example, the N-terminal position can contain a positively charged amino acid, such as lysine, arginine, or histidine.

Typically, a residue of a polypeptide is conserved when a class of amino acids or a single amino acid is found at a particular position in at least about 40% of all class members; more typically, at least about 50%; even more typically, at least about 60% of the members. Usually, a residue is conserved when a class or single amino acid is found in at least about 70% of the members of a family or motif; more usually, at least about 80%; even more usually, at least about 90%; even more usually, at least about 95%.

A residue is considered conserved when three unrelated amino acids are found at a particular position in the some or all of the members; more usually, two unrelated amino acids. These residues are conserved when the unrelated amino acids are found at particular positions in at least about 40% of all class member; more typically, at least about 50%; even more typically, at least about 60% of the members. Usually, a residue is conserved when a class or single amino acid is found in at least about 70% of the members of a family or motif; more usually, at least about 80%; even more usually, at least about 90%; even more usually, at least about 95%.

A query sequence has similarity to a profile or MSA when the query sequence comprises at least about 25% of the conserved residues of the profile or MSA; more usually, at least about 30%; even more usually; at least about 40%. Typically, the query sequence has a stronger similarity to a profile sequence or MSA when the query sequence comprises at least about 45% of the conserved residues of the profile or MSA; more typically, at least about 50%; even more typically; at least about 55%.

Identification of Secreted & Membrane-Bound Polypeptides

Both secreted and membrane-bound polypeptides of the present invention are of particular interest. For example, levels of secreted polypeptides can be assayed in body fluids that are convenient, such as blood, plasma, serum, and other body fluids such as urine, prostatic fluid and semen. Membrane-bound polypeptides are useful for constructing vaccine antigens or inducing an immune response. Such antigens would comprise all or part of the extracellular region of the membrane-bound polypeptides. Because both secreted and membrane-bound polypeptides comprise a fragment of contiguous hydrophobic amino acids, hydrophobicity predicting algorithms can be used to identify such polypeptides.

A signal sequence is usually encoded by both secreted and membrane-bound polypeptide genes to direct a polypeptide to the surface of the cell. The signal sequence usually comprises a stretch of hydrophobic residues. Such signal sequences can fold into helical structures. Membrane-bound polypeptides typically comprise at least one transmembrane region that possesses a stretch of hydrophobic amino acids that can transverse the membrane. Some transmembrane regions also exhibit a helical structure. Hydrophobic fragments within a polypeptide can be identified by using computer algorithms. Such algorithms include Hopp & Woods, *Proc. Natl. Acad. Sci. USA* (1981) 78:3824-3828; Kyte & Doolittle, *J. Mol. Biol.* (1982) 157: 105-132; and RAOAR algorithm, Degli Esposti *et al.*, *Eur. J. Biochem.* (1990) 190: 207-219.

Another method of identifying secreted and membrane-bound polypeptides is to translate the polynucleotides of the invention in all six frames and determine if at least 8 contiguous hydrophobic amino acids are present. Those translated polypeptides with at least 8; more typically, 10; even more typically, 12 contiguous hydrophobic amino acids are considered to be either a putative secreted or membrane bound polypeptide. Hydrophobic amino acids include alanine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, threonine, tryptophan, tyrosine, and valine.

Identification of the Function of an Expression Product of a Full-Length Gene

Ribozymes, antisense constructs, and dominant negative mutants can be used to determine function of the expression product of a gene corresponding to a polynucleotide provided herein.

These methods and compositions are particularly useful where the provided novel polynucleotide exhibits no significant or substantial homology to a sequence encoding a gene of known function. Antisense molecules and ribozymes can be constructed from synthetic polynucleotides. Typically, the phosphoramidite method of oligonucleotide synthesis is used. See Beaucage *et al.*, *Tet. Lett.* (1981) 22:1859 and USPN 4,668,777. Automated devices for synthesis are available to create oligonucleotides using this chemistry. Examples of such devices include Biosearch 8600, Models 392 and 394 by Applied Biosystems, a division of Perkin-Elmer Corp., Foster City, California, USA; and Expedite by Perceptive Biosystems, Framingham, Massachusetts, USA. Synthetic RNA, phosphate analog oligonucleotides, and chemically derivatized oligonucleotides can also be produced, and can be covalently attached to other molecules. RNA oligonucleotides can be synthesized, for example, using RNA phosphoramidites. This method can be performed on an automated synthesizer, such as Applied Biosystems, Models 392 and 394, Foster City, California, USA.

Phosphorothioate oligonucleotides can also be synthesized for antisense construction. A sulfurizing reagent, such as tetraethylthiuram disulfide (TETD) in acetonitrile can be used to convert the internucleotide cyanoethyl phosphite to the phosphorothioate triester within 15 minutes at room temperature. TETD replaces the iodine reagent, while all other reagents used for standard

phosphoramidite chemistry remain the same. Such a synthesis method can be automated using Models 392 and 394 by Applied Biosystems, for example.

Oligonucleotides of up to 200 nt can be synthesized, more typically, 100 nt, more typically 50 nt; even more typically 30 to 40 nt. These synthetic fragments can be annealed and ligated together to construct larger fragments. See, for example, Sambrook *et al.*, *supra*. Trans-cleaving catalytic RNAs (ribozymes) are RNA molecules possessing endoribonuclease activity. Ribozymes are specifically designed for a particular target, and the target message must contain a specific nucleotide sequence. They are engineered to cleave any RNA species site-specifically in the background of cellular RNA. The cleavage event renders the mRNA unstable and prevents protein expression. Importantly, ribozymes can be used to inhibit expression of a gene of unknown function for the purpose of determining its function in an in vitro or in vivo context, by detecting the phenotypic effect. One commonly used ribozyme motif is the hammerhead, for which the substrate sequence requirements are minimal. Design of the hammerhead ribozyme, as well as therapeutic uses of ribozymes, are disclosed in Usman *et al.*, *Current Opin. Struct. Biol.* (1996) 6:527. Methods for production of ribozymes, including hairpin structure ribozyme fragments, methods of increasing ribozyme specificity, and the like are known in the art.

The hybridizing region of the ribozyme can be modified or can be prepared as a branched structure as described in Horn and Urdea, *Nucleic Acids Res.* (1989) 17:6959. The basic structure of the ribozymes can also be chemically altered in ways familiar to those skilled in the art, and chemically synthesized ribozymes can be administered as synthetic oligonucleotide derivatives modified by monomeric units. In a therapeutic context, liposome mediated delivery of ribozymes improves cellular uptake, as described in Birikh *et al.*, *Eur. J. Biochem.* (1997) 245:1.

Antisense nucleic acids are designed to specifically bind to RNA, resulting in the formation of RNA-DNA or RNA-RNA hybrids, with an arrest of DNA replication, reverse transcription or messenger RNA translation. Antisense polynucleotides based on a selected polynucleotide sequence can interfere with expression of the corresponding gene. Antisense polynucleotides are typically generated within the cell by expression from antisense constructs that contain the antisense strand as the transcribed strand. Antisense polynucleotides based on the disclosed polynucleotides will bind and/or interfere with the translation of mRNA comprising a sequence complementary to the antisense polynucleotide. The expression products of control cells and cells treated with the antisense construct are compared to detect the protein product of the gene corresponding to the polynucleotide upon which the antisense construct is based. The protein is isolated and identified using routine biochemical methods.

Given the extensive background literature and clinical experience in antisense therapy, one skilled in the art can use selected polynucleotides of the invention as additional potential therapeutics. The choice of polynucleotide can be narrowed by first testing them for binding to

"hot spot" regions of the genome of cancerous cells. If a polynucleotide is identified as binding to a "hot spot", testing the polynucleotide as an antisense compound in the corresponding cancer cells is warranted.

As an alternative method for identifying function of the gene corresponding to a polynucleotide disclosed herein, dominant negative mutations are readily generated for corresponding proteins that are active as homomultimers. A mutant polypeptide will interact with wild-type polypeptides (made from the other allele) and form a non-functional multimer. Thus, a mutation is in a substrate-binding domain, a catalytic domain, or a cellular localization domain. Preferably, the mutant polypeptide will be overproduced. Point mutations are made that have such an effect. In addition, fusion of different polypeptides of various lengths to the terminus of a protein can yield dominant negative mutants. General strategies are available for making dominant negative mutants (see, *e.g.*, Herskowitz, *Nature* (1987) 329:219). Such techniques can be used to create loss of function mutations, which are useful for determining protein function.

Polypeptides and Variants Thereof

The polypeptides of the invention include those encoded by the disclosed polynucleotides, as well as nucleic acids that, by virtue of the degeneracy of the genetic code, are not identical in sequence to the disclosed polynucleotides. Thus, the invention includes within its scope a polypeptide encoded by a polynucleotide having the sequence of any one of SEQ ID NOS:1-316 or a variant thereof.

In general, the term "polypeptide" as used herein refers to both the full length polypeptide encoded by the recited polynucleotide, the polypeptide encoded by the gene represented by the recited polynucleotide, as well as portions or fragments thereof. "Polypeptides" also includes variants of the naturally occurring proteins, where such variants are homologous or substantially similar to the naturally occurring protein, and can be of an origin of the same or different species as the naturally occurring protein (*e.g.*, human, murine, or some other species that naturally expresses the recited polypeptide, usually a mammalian species). In general, variant polypeptides have a sequence that has at least about 80%, usually at least about 90%, and more usually at least about 98% sequence identity with a differentially expressed polypeptide of the invention, as measured by BLAST 2.0 using the parameters described above. The variant polypeptides can be naturally or non-naturally glycosylated, *i.e.*, the polypeptide has a glycosylation pattern that differs from the glycosylation pattern found in the corresponding naturally occurring protein.

The invention also encompasses homologs of the disclosed polypeptides (or fragments thereof) where the homologs are isolated from other species, *i.e.* other animal or plant species, where such homologs, usually mammalian species, *e.g.* rodents, such as mice, rats; domestic animals, *e.g.*, horse, cow, dog, cat; and humans. By "homolog" is meant a polypeptide having at least about 35%, usually at least about 40% and more usually at least about 60% amino acid

sequence identity to a particular differentially expressed protein as identified above, where sequence identity is determined using the BLAST 2.0 algorithm, with the parameters described *supra*.

In general, the polypeptides of the subject invention are provided in a non-naturally occurring environment, *e.g.* are separated from their naturally occurring environment. In certain embodiments, the subject protein is present in a composition that is enriched for the protein as compared to a control. As such, purified polypeptide is provided, where by purified is meant that the protein is present in a composition that is substantially free of non-differentially expressed polypeptides, where by substantially free is meant that less than 90%, usually less than 60% and more usually less than 50% of the composition is made up of non-differentially expressed polypeptides.

Also within the scope of the invention are variants; variants of polypeptides include mutants, fragments, and fusions. Mutants can include amino acid substitutions, additions or deletions. The amino acid substitutions can be conservative amino acid substitutions or substitutions to eliminate non-essential amino acids, such as to alter a glycosylation site, a phosphorylation site or an acetylation site, or to minimize misfolding by substitution or deletion of one or more cysteine residues that are not necessary for function. Conservative amino acid substitutions are those that preserve the general charge, hydrophobicity/ hydrophilicity, and/or steric bulk of the amino acid substituted. Variants can be designed so as to retain or have enhanced biological activity of a particular region of the protein (*e.g.*, a functional domain and/or, where the polypeptide is a member of a protein family, a region associated with a consensus sequence). Selection of amino acid alterations for production of variants can be based upon the accessibility (interior vs. exterior) of the amino acid (see, *e.g.*, Go *et al.*, *Int. J. Peptide Protein Res.* (1980) 15:211), the thermostability of the variant polypeptide (see, *e.g.*, Querol *et al.*, *Prot. Eng.* (1996) 9:265), desired glycosylation sites (see, *e.g.*, Olsen and Thomsen, *J. Gen. Microbiol.* (1991) 137:579), desired disulfide bridges (see, *e.g.*, Clarke *et al.*, *Biochemistry* (1993) 32:4322; and Wakarchuk *et al.*, *Protein Eng.* (1994) 7:1379), desired metal binding sites (see, *e.g.*, Toma *et al.*, *Biochemistry* (1991) 30:97, and Haezebrouck *et al.*, *Protein Eng.* (1993) 6:643), and desired substitutions with in proline loops (see, *e.g.*, Masul *et al.*, *Appl. Env. Microbiol.* (1994) 60:3579). Cysteine-depleted muteins can be produced as disclosed in USPN 4,959,314.

Variants also include fragments of the polypeptides disclosed herein, particularly biologically active fragments and/or fragments corresponding to functional domains. Fragments of interest will typically be at least about 10 aa to at least about 15 aa in length, usually at least about 50 aa in length, and can be as long as 300 aa in length or longer, but will usually not exceed about 1000 aa in length, where the fragment will have a stretch of amino acids that is identical to a polypeptide encoded by a polynucleotide having a sequence of any SEQ ID NOS:1-316, or a

homolog thereof. The protein variants described herein are encoded by polynucleotides that are within the scope of the invention. The genetic code can be used to select the appropriate codons to construct the corresponding variants.

Computer-Related Embodiments

5 In general, a library of polynucleotides is a collection of sequence information, which information is provided in either biochemical form (*e.g.*, as a collection of polynucleotide molecules), or in electronic form (*e.g.*, as a collection of polynucleotide sequences stored in a computer-readable form, as in a computer system and/or as part of a computer program). The sequence information of the polynucleotides can be used in a variety of ways, *e.g.*, as a resource for
10 gene discovery, as a representation of sequences expressed in a selected cell type (*e.g.*, cell type markers), and/or as markers of a given disease or disease state. In general, a disease marker is a representation of a gene product that is present in all cells affected by disease either at an increased or decreased level relative to a normal cell (*e.g.*, a cell of the same or similar type that is not substantially affected by disease). For example, a polynucleotide sequence in a library can be a
15 polynucleotide that represents an mRNA, polypeptide, or other gene product encoded by the polynucleotide, that is either overexpressed or underexpressed in a breast ductal cell affected by cancer relative to a normal (*i.e.*, substantially disease-free) breast cell.

The nucleotide sequence information of the library can be embodied in any suitable form, *e.g.*, electronic or biochemical forms. For example, a library of sequence information embodied in
20 electronic form comprises an accessible computer data file (or, in biochemical form, a collection of nucleic acid molecules) that contains the representative nucleotide sequences of genes that are differentially expressed (*e.g.*, overexpressed or underexpressed) as between, for example, i) a cancerous cell and a normal cell; ii) a cancerous cell and a dysplastic cell; iii) a cancerous cell and a cell affected by a disease or condition other than cancer; iv) a metastatic cancerous cell and a
25 normal cell and/or non-metastatic cancerous cell; v) a malignant cancerous cell and a non-malignant cancerous cell (or a normal cell) and/or vi) a dysplastic cell relative to a normal cell. Other combinations and comparisons of cells affected by various diseases or stages of disease will be readily apparent to the ordinarily skilled artisan. Biochemical embodiments of the library include a collection of nucleic acids that have the sequences of the genes in the library, where the
30 nucleic acids can correspond to the entire gene in the library or to a fragment thereof, as described in greater detail below.

The polynucleotide libraries of the subject invention generally comprise sequence information of a plurality of polynucleotide sequences, where at least one of the polynucleotides has a sequence of any of SEQ ID NOS:1-316. By plurality is meant at least 2, usually at least 3
35 and can include up to all of SEQ ID NOS:1-316. The length and number of polynucleotides in the

library will vary with the nature of the library, *e.g.*, if the library is an oligonucleotide array, a cDNA array, a computer database of the sequence information, etc.

Where the library is an electronic library, the nucleic acid sequence information can be present in a variety of media. "Media" refers to a manufacture, other than an isolated nucleic acid molecule, that contains the sequence information of the present invention. Such a manufacture provides the genome sequence or a subset thereof in a form that can be examined by means not directly applicable to the sequence as it exists in a nucleic acid. For example, the nucleotide sequence of the present invention, *e.g.* the nucleic acid sequences of any of the polynucleotides of SEQ ID NOS:1-316, can be recorded on computer readable media, *e.g.* any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as a floppy disc, a hard disc storage medium, and a magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. One of skill in the art can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising a recording of the present sequence information. "Recorded" refers to a process for storing information on computer readable medium, using any such methods as known in the art. Any convenient data storage structure can be chosen, based on the means used to access the stored information. A variety of data processor programs and formats can be used for storage, *e.g.* word processing text file, database format, *etc.* In addition to the sequence information, electronic versions of the libraries of the invention can be provided in conjunction or connection with other computer-readable information and/or other types of computer-readable files (*e.g.*, searchable files, executable files, *etc.*, including, but not limited to, for example, search program software, *etc.*).

By providing the nucleotide sequence in computer readable form, the information can be accessed for a variety of purposes. Computer software to access sequence information is publicly available. For example, the gapped BLAST (Altschul *et al. Nucleic Acids Res.* (1997) 25:3389-3402) and BLAZE (Brutlag *et al. Comp. Chem.* (1993) 17:203) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) within the genome that contain homology to ORFs from other organisms.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based system are suitable for use in the present invention. The data storage means can comprise any manufacture comprising a recording of the present sequence information as described above, or a memory access means that can access such a manufacture.

"Search means" refers to one or more programs implemented on the computer-based system, to compare a target sequence or target structural motif, or expression levels of a polynucleotide in a sample, with the stored sequence information. Search means can be used to identify fragments or regions of the genome that match a particular target sequence or target motif.

5 A variety of known algorithms are publicly known and commercially available, *e.g.* MacPattern (EMBL), BLASTN and BLASTX (NCBI). A "target sequence" can be any polynucleotide or amino acid sequence of six or more contiguous nucleotides or two or more amino acids, preferably from about 10 to 100 amino acids or from about 30 to 300 nt. A variety of comparing means can be used to accomplish comparison of sequence information from a sample (*e.g.*, to analyze target
10 sequences, target motifs, or relative expression levels) with the data storage means. A skilled artisan can readily recognize that any one of the publicly available homology search programs can be used as the search means for the computer based systems of the present invention to accomplish comparison of target sequences and motifs. Computer programs to analyze expression levels in a sample and in controls are also known in the art.

15 A "target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration that is formed upon the folding of the target motif, or on consensus sequences of regulatory or active sites. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target
20 motifs include, but are not limited to, hairpin structures, promoter sequences and other expression elements such as binding sites for transcription factors.

A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the present invention. One format for an output means ranks the relative expression levels of different polynucleotides. Such presentation
25 provides a skilled artisan with a ranking of relative expression levels to determine a gene expression profile.

As discussed above, the "library" of the invention also encompasses biochemical libraries of the polynucleotides of SEQ ID NOS:1-316, *e.g.*, collections of nucleic acids representing the provided polynucleotides. The biochemical libraries can take a variety of forms, *e.g.*, a solution of
30 cDNAs, a pattern of probe nucleic acids stably associated with a surface of a solid support (*i.e.*, an array) and the like. Of particular interest are nucleic acid arrays in which one or more of SEQ ID NOS:1-316 is represented on the array. By array is meant an article of manufacture that has at least a substrate with at least two distinct nucleic acid targets on one of its surfaces, where the number of distinct nucleic acids can be considerably higher, typically being at least 10 nt, usually at
35 least 20 nt and often at least 25 nt. A variety of different array formats have been developed and are known to those of skill in the art. The arrays of the subject invention find use in a variety of

applications, including gene expression analysis, drug screening, mutation analysis and the like, as disclosed in the above-listed exemplary patent documents.

In addition to the above nucleic acid libraries, analogous libraries of polypeptides are also provided, where the where the polypeptides of the library will represent at least a portion of the polypeptides encoded by SEQ ID NOS:1-316.

Utilities

Use of Polynucleotide Probes in Mapping, and in Tissue Profiling

Polynucleotide probes, generally comprising at least 12 contiguous nt of a polynucleotide as shown in the Sequence Listing, are used for a variety of purposes, such as chromosome mapping of the polynucleotide and detection of transcription levels. Additional disclosure about preferred regions of the disclosed polynucleotide sequences is found in the Examples. A probe that hybridizes specifically to a polynucleotide disclosed herein should provide a detection signal at least 5-, 10-, or 20-fold higher than the background hybridization provided with other unrelated sequences.

Detection of Expression Levels. Nucleotide probes are used to detect expression of a gene corresponding to the provided polynucleotide. In Northern blots, mRNA is separated electrophoretically and contacted with a probe. A probe is detected as hybridizing to an mRNA species of a particular size. The amount of hybridization is quantitated to determine relative amounts of expression, for example under a particular condition. Probes are used for in situ hybridization to cells to detect expression. Probes can also be used *in vivo* for diagnostic detection of hybridizing sequences. Probes are typically labeled with a radioactive isotope. Other types of detectable labels can be used such as chromophores, fluors, and enzymes. Other examples of nucleotide hybridization assays are described in WO92/02526 and USPN 5,124,246.

Alternatively, the Polymerase Chain Reaction (PCR) is another means for detecting small amounts of target nucleic acids (see, e.g., Mullis *et al.*, *Meth. Enzymol.* (1987) 155:335; USPN 4,683,195; and USPN 4,683,202). Two primer polynucleotides nucleotides that hybridize with the target nucleic acids are used to prime the reaction. The primers can be composed of sequence within or 3' and 5' to the polynucleotides of the Sequence Listing. Alternatively, if the primers are 3' and 5' to these polynucleotides, they need not hybridize to them or the complements. After amplification of the target with a thermostable polymerase, the amplified target nucleic acids can be detected by methods known in the art, e.g., Southern blot. mRNA or cDNA can also be detected by traditional blotting techniques (e.g., Southern blot, Northern blot, etc.) described in Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual" (New York, Cold Spring Harbor Laboratory, 1989) (e.g., without PCR amplification). In general, mRNA or cDNA generated from mRNA using a polymerase enzyme can be purified and separated using gel electrophoresis, and transferred to a

solid support, such as nitrocellulose. The solid support is exposed to a labeled probe, washed to remove any unhybridized probe, and duplexes containing the labeled probe are detected.

Mapping. Polynucleotides of the present invention can be used to identify a chromosome on which the corresponding gene resides. Such mapping can be useful in identifying the function of the polynucleotide-related gene by its proximity to other genes with known function. Function can also be assigned to the polynucleotide-related gene when particular syndromes or diseases map to the same chromosome. For example, use of polynucleotide probes in identification and quantification of nucleic acid sequence aberrations is described in USPN 5,783,387. An exemplary mapping method is fluorescence in situ hybridization (FISH), which facilitates comparative genomic hybridization to allow total genome assessment of changes in relative copy number of DNA sequences (see, e.g., Valdes *et al.*, *Methods in Molecular Biology* (1997) 68:1). Polynucleotides can also be mapped to particular chromosomes using, for example, radiation hybrids or chromosome-specific hybrid panels. See Leach *et al.*, *Advances in Genetics*, (1995) 33:63-99; Walter *et al.*, *Nature Genetics* (1994) 7:22; Walter and Goodfellow, *Trends in Genetics* (1992) 9:352. Panels for radiation hybrid mapping are available from Research Genetics, Inc., Huntsville, Alabama, USA. Databases for markers using various panels are available via the world wide web at sites supported by the Stanford Human Genome Center (Stanford University) and the Whitehead Institute for Biomedical Research/MIT Center for Genome Research. The statistical program RHMAP can be used to construct a map based on the data from radiation hybridization with a measure of the relative likelihood of one order versus another. RHMAP is available via the world wide web at a site supported by the Center for Statistical Genetics at the University of Michigan School of Public Health. In addition, commercial programs are available for identifying regions of chromosomes commonly associated with disease, such as cancer.

Tissue Typing or Profiling. Expression of specific mRNA corresponding to the provided polynucleotides can vary in different cell types and can be tissue-specific. This variation of mRNA levels in different cell types can be exploited with nucleic acid probe assays to determine tissue types. For example, PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes substantially identical or complementary to polynucleotides listed in the Sequence Listing can determine the presence or absence of the corresponding cDNA or mRNA.

Tissue typing can be used to identify the developmental organ or tissue source of a metastatic lesion by identifying the expression of a particular marker of that organ or tissue. If a polynucleotide is expressed only in a specific tissue type, and a metastatic lesion is found to express that polynucleotide, then the developmental source of the lesion has been identified. Expression of a particular polynucleotide can be assayed by detection of either the corresponding mRNA or the protein product. As would be readily apparent to any forensic scientist, the sequences disclosed herein are useful in differentiating human tissue from non-human tissue. In particular, these

sequences are useful to differentiate human tissue from bird, reptile, and amphibian tissue, for example.

Use of Polymorphisms. A polynucleotide of the invention can be used in forensics, genetic analysis, mapping, and diagnostic applications where the corresponding region of a gene is polymorphic in the human population. Any means for detecting a polymorphism in a gene can be used, including, but not limited to electrophoresis of protein polymorphic variants, differential sensitivity to restriction enzyme cleavage, and hybridization to allele-specific probes.

Antibody Production

Expression products of a polynucleotide of the invention, as well as the corresponding mRNA, cDNA, or complete gene, can be prepared and used for raising antibodies for experimental, diagnostic, and therapeutic purposes. For polynucleotides to which a corresponding gene has not been assigned, this provides an additional method of identifying the corresponding gene. The polynucleotide or related cDNA is expressed as described above, and antibodies are prepared. These antibodies are specific to an epitope on the polypeptide encoded by the polynucleotide, and can precipitate or bind to the corresponding native protein in a cell or tissue preparation or in a cell-free extract of an in vitro expression system.

Methods for production of antibodies that specifically bind a selected antigen are well known in the art. Immunogens for raising antibodies can be prepared by mixing a polypeptide encoded by a polynucleotide of the invention with an adjuvant, and/or by making fusion proteins with larger immunogenic proteins. Polypeptides can also be covalently linked to other larger immunogenic proteins, such as keyhole limpet hemocyanin. Immunogens are typically administered intradermally, subcutaneously, or intramuscularly to experimental animals such as rabbits, sheep, and mice, to generate antibodies. Monoclonal antibodies can be generated by isolating spleen cells and fusing myeloma cells to form hybridomas. Alternatively, the selected polynucleotide is administered directly, such as by intramuscular injection, and expressed in vivo. The expressed protein generates a variety of protein-specific immune responses, including production of antibodies, comparable to administration of the protein.

Preparations of polyclonal and monoclonal antibodies specific for polypeptides encoded by a selected polynucleotide are made using standard methods known in the art. The antibodies specifically bind to epitopes present in the polypeptides encoded by polynucleotides disclosed in the Sequence Listing. Typically, at least 6, 8, 10, or 12 contiguous amino acids are required to form an epitope. Epitopes that involve non-contiguous amino acids may require a longer polypeptide, e.g., at least 15, 25, or 50 amino acids. Antibodies that specifically bind to human polypeptides encoded by the provided polypeptides should provide a detection signal at least 5-, 10-, or 20-fold higher than a detection signal provided with other proteins when used in Western blots or other immunochemical assays. Preferably, antibodies that specifically bind to polypeptides of the

invention do not bind to other proteins in immunochemical assays at detectable levels and can immunoprecipitate the specific polypeptide from solution.

The invention also contemplates naturally occurring antibodies specific for a polypeptide of the invention. For example, serum antibodies to a polypeptide of the invention in a human population can be purified by methods well known in the art, e.g., by passing antiserum over a column to which the corresponding selected polypeptide or fusion protein is bound. The bound antibodies can then be eluted from the column, for example using a buffer with a high salt concentration.

In addition to the antibodies discussed above, the invention also contemplates genetically engineered antibodies, antibody derivatives (e.g., single chain antibodies, antibody fragments (e.g., Fab, etc.)), according to methods well known in the art.

Polynucleotides or Arrays for Diagnostics

Polynucleotide arrays provide a high throughput technique that can assay a large number of polynucleotide sequences in a sample. This technology can be used as a diagnostic and as a tool to test for differential expression, e.g., to determine function of an encoded protein. Arrays can be created by spotting polynucleotide probes onto a substrate (e.g., glass, nitrocellulose, etc.) in a two-dimensional matrix or array having bound probes. The probes can be bound to the substrate by either covalent bonds or by non-specific interactions, such as hydrophobic interactions. Samples of polynucleotides can be detectably labeled (e.g., using radioactive or fluorescent labels) and then hybridized to the probes. Double stranded polynucleotides, comprising the labeled sample polynucleotides bound to probe polynucleotides, can be detected once the unbound portion of the sample is washed away. Techniques for constructing arrays and methods of using these arrays are described in EP 799 897; WO 97/29212; WO 97/27317; EP 785 280; WO 97/02357; USPN 5,593,839; USPN 5,578,832; EP 728 520; USPN 5,599,695; EP 721 016; USPN 5,556,752; WO 95/22058; and USPN 5,631,734. Arrays can be used to, for example, examine differential expression of genes and can be used to determine gene function. For example, arrays can be used to detect differential expression of a polynucleotide between a test cell and control cell (e.g., cancer cells and normal cells). For example, high expression of a particular message in a cancer cell, which is not observed in a corresponding normal cell, can indicate a cancer specific gene product. Exemplary uses of arrays are further described in, for example, Pappalarado *et al.*, *Sem. Radiation Oncol.* (1998) 8:217; and Ramsay *Nature Biotechnol.* (1998) 16:40.

Differential Expression in Diagnosis

The polynucleotides of the invention can also be used to detect differences in expression levels between two cells, e.g., as a method to identify abnormal or diseased tissue in a human. For polynucleotides corresponding to profiles of protein families, the choice of tissue can be selected according to the putative biological function. In general, the expression of a gene corresponding to

a specific polynucleotide is compared between a first tissue that is suspected of being diseased and a second, normal tissue of the human. The tissue suspected of being abnormal or diseased can be derived from a different tissue type of the human, but preferably it is derived from the same tissue type; for example an intestinal polyp or other abnormal growth should be compared with normal intestinal tissue. The normal tissue can be the same tissue as that of the test sample, or any normal tissue of the patient, especially those that express the polynucleotide-related gene of interest (*e.g.*, brain, thymus, testis, heart, prostate, placenta, spleen, small intestine, skeletal muscle, pancreas, and the mucosal lining of the colon). A difference between the polynucleotide-related gene, mRNA, or protein in the two tissues which are compared, for example in molecular weight, amino acid or nucleotide sequence, or relative abundance, indicates a change in the gene, or a gene which regulates it, in the tissue of the human that was suspected of being diseased. Examples of detection of differential expression and its use in diagnosis of cancer are described in USPNs 5,688,641 and 5,677,125.

A genetic predisposition to disease in a human can also be detected by comparing expression levels of an mRNA or protein corresponding to a polynucleotide of the invention in a fetal tissue with levels associated in normal fetal tissue. Fetal tissues that are used for this purpose include, but are not limited to, amniotic fluid, chorionic villi, blood, and the blastomere of an in vitro-fertilized embryo. The comparable normal polynucleotide-related gene is obtained from any tissue. The mRNA or protein is obtained from a normal tissue of a human in which the polynucleotide-related gene is expressed. Differences such as alterations in the nucleotide sequence or size of the same product of the fetal polynucleotide-related gene or mRNA, or alterations in the molecular weight, amino acid sequence, or relative abundance of fetal protein, can indicate a germline mutation in the polynucleotide-related gene of the fetus, which indicates a genetic predisposition to disease. In general, diagnostic, prognostic, and other methods of the invention based on differential expression involve detection of a level or amount of a gene product, particularly a differentially expressed gene product, in a test sample obtained from a patient suspected of having or being susceptible to a disease (*e.g.*, breast cancer, lung cancer, colon cancer and/or metastatic forms thereof), and comparing the detected levels to those levels found in normal cells (*e.g.*, cells substantially unaffected by cancer) and/or other control cells (*e.g.*, to differentiate a cancerous cell from a cell affected by dysplasia). Furthermore, the severity of the disease can be assessed by comparing the detected levels of a differentially expressed gene product with those levels detected in samples representing the levels of differentially gene product associated with varying degrees of severity of disease. It should be noted that use of the term “diagnostic” herein is not necessarily meant to exclude “prognostic” or “prognosis,” but rather is used as a matter of convenience.

The term "differentially expressed gene" is generally intended to encompass a polynucleotide that can, for example, include an open reading frame encoding a gene product (*e.g.*, a polypeptide), and/or introns of such genes and adjacent 5' and 3' non-coding nucleotide sequences involved in the regulation of expression, up to about 20 kb beyond the coding region, but possibly further in either direction. The gene can be introduced into an appropriate vector for extrachromosomal maintenance or for integration into a host genome. In general, a difference in expression level associated with a decrease in expression level of at least about 25%, usually at least about 50% to 75%, more usually at least about 90% or more is indicative of a differentially expressed gene of interest, *i.e.*, a gene that is underexpressed or down-regulated in the test sample relative to a control sample. Furthermore, a difference in expression level associated with an increase in expression of at least about 25%, usually at least about 50% to 75%, more usually at least about 90% and can be at least about 1 ½-fold, usually at least about 2-fold to about 10-fold, and can be about 100-fold to about 1,000-fold increase relative to a control sample is indicative of a differentially expressed gene of interest, *i.e.*, an overexpressed or up-regulated gene.

"Differentially expressed polynucleotide" as used herein means a nucleic acid molecule (RNA or DNA) comprising a sequence that represents a differentially expressed gene, *e.g.*, the differentially expressed polynucleotide comprises a sequence (*e.g.*, an open reading frame encoding a gene product) that uniquely identifies a differentially expressed gene so that detection of the differentially expressed polynucleotide in a sample is correlated with the presence of a differentially expressed gene in a sample. "Differentially expressed polynucleotides" is also meant to encompass fragments of the disclosed polynucleotides, *e.g.*, fragments retaining biological activity, as well as nucleic acids homologous, substantially similar, or substantially identical (*e.g.*, having about 90% sequence identity) to the disclosed polynucleotides.

"Diagnosis" as used herein generally includes determination of a subject's susceptibility to a disease or disorder, determination as to whether a subject is presently affected by a disease or disorder, as well as to the prognosis of a subject affected by a disease or disorder (*e.g.*, identification of pre-metastatic or metastatic cancerous states, stages of cancer, or responsiveness of cancer to therapy). The present invention particularly encompasses diagnosis of subjects in the context of breast cancer (*e.g.*, carcinoma in situ (*e.g.*, ductal carcinoma in situ), estrogen receptor (ER)-positive breast cancer, ER-negative breast cancer, or other forms and/or stages of breast cancer), lung cancer (*e.g.*, small cell carcinoma, non-small cell carcinoma, mesothelioma, and other forms and/or stages of lung cancer), and colon cancer (*e.g.*, adenomatous polyp, colorectal carcinoma, and other forms and/or stages of colon cancer).

"Sample" or "biological sample" as used throughout here are generally meant to refer to samples of biological fluids or tissues, particularly samples obtained from tissues, especially from cells of the type associated with the disease for which the diagnostic application is designed (*e.g.*,

ductal adenocarcinoma), and the like. "Samples" is also meant to encompass derivatives and fractions of such samples (*e.g.*, cell lysates). Where the sample is solid tissue, the cells of the tissue can be dissociated or tissue sections can be analyzed.

Methods of the subject invention useful in diagnosis or prognosis typically involve comparison of the abundance of a selected differentially expressed gene product in a sample of interest with that of a control to determine any relative differences in the expression of the gene product, where the difference can be measured qualitatively and/or quantitatively. Quantitation can be accomplished, for example, by comparing the level of expression product detected in the sample with the amounts of product present in a standard curve. A comparison can be made visually; by using a technique such as densitometry, with or without computerized assistance; by preparing a representative library of cDNA clones of mRNA isolated from a test sample, sequencing the clones in the library to determine that number of cDNA clones corresponding to the same gene product, and analyzing the number of clones corresponding to that same gene product relative to the number of clones of the same gene product in a control sample; or by using an array to detect relative levels of hybridization to a selected sequence or set of sequences, and comparing the hybridization pattern to that of a control. The differences in expression are then correlated with the presence or absence of an abnormal expression pattern. A variety of different methods for determining the nucleic acid abundance in a sample are known to those of skill in the art (see, *e.g.*, WO 97/27317).

In general, diagnostic assays of the invention involve detection of a gene product of a the polynucleotide sequence (*e.g.*, mRNA or polypeptide) that corresponds to a sequence of SEQ ID NOS:1-316. The patient from whom the sample is obtained can be apparently healthy, susceptible to disease (*e.g.*, as determined by family history or exposure to certain environmental factors), or can already be identified as having a condition in which altered expression of a gene product of the invention is implicated.

Diagnosis can be determined based on detected gene product expression levels of a gene product encoded by at least one, preferably at least two or more, at least 3 or more, or at least 4 or more of the polynucleotides having a sequence set forth in SEQ ID NOS:1-316, and can involve detection of expression of genes corresponding to all of SEQ ID NOS:1-316 and/or additional sequences that can serve as additional diagnostic markers and/or reference sequences. Where the diagnostic method is designed to detect the presence or susceptibility of a patient to cancer, the assay preferably involves detection of a gene product encoded by a gene corresponding to a polynucleotide that is differentially expressed in cancer. Examples of such differentially expressed polynucleotides are described in the Examples below. Given the provided polynucleotides and information regarding their relative expression levels provided herein, assays using such polynucleotides and detection of their expression levels in diagnosis and prognosis will be readily apparent to the ordinarily skilled artisan.

Any of a variety of detectable labels can be used in connection with the various embodiments of the diagnostic methods of the invention. Suitable detectable labels include fluorochromes, (e.g. fluorescein isothiocyanate (FITC), rhodamine, Texas Red, phycoerythrin, allophycocyanin, 6-carboxyfluorescein (6-FAM), 2',7'-dimethoxy-4',5'-dichloro-6-carboxyfluorescein, 6-carboxy-X-rhodamine (ROX), 6-carboxy-2',4',7',4,7-hexachlorofluorescein (HEX), 5-carboxyfluorescein (5-FAM) or N,N,N',N'-tetramethyl-6-carboxyrhodamine (TAMRA)), radioactive labels, (e.g. ^{32}P , ^{35}S , ^3H , *etc.*), and the like. The detectable label can involve a two stage systems (e.g., biotin-avidin, hapten-anti-hapten antibody, *etc.*)

Reagents specific for the polynucleotides and polypeptides of the invention, such as antibodies and nucleotide probes, can be supplied in a kit for detecting the presence of an expression product in a biological sample. The kit can also contain buffers or labeling components, as well as instructions for using the reagents to detect and quantify expression products in the biological sample. Exemplary embodiments of the diagnostic methods of the invention are described below in more detail.

Polypeptide detection in diagnosis. In one embodiment, the test sample is assayed for the level of a differentially expressed polypeptide. Diagnosis can be accomplished using any of a number of methods to determine the absence or presence or altered amounts of the differentially expressed polypeptide in the test sample. For example, detection can utilize staining of cells or histological sections with labeled antibodies, performed in accordance with conventional methods. Cells can be permeabilized to stain cytoplasmic molecules. In general, antibodies that specifically bind a differentially expressed polypeptide of the invention are added to a sample, and incubated for a period of time sufficient to allow binding to the epitope, usually at least about 10 minutes. The antibody can be detectably labeled for direct detection (e.g., using radioisotopes, enzymes, fluorescers, chemiluminescers, and the like), or can be used in conjunction with a second stage antibody or reagent to detect binding (e.g., biotin with horseradish peroxidase-conjugated avidin, a secondary antibody conjugated to a fluorescent compound, e.g. fluorescein, rhodamine, Texas red, *etc.*). The absence or presence of antibody binding can be determined by various methods, including flow cytometry of dissociated cells, microscopy, radiography, scintillation counting, *etc.* Any suitable alternative methods can of qualitative or quantitative detection of levels or amounts of differentially expressed polypeptide can be used, for example ELISA, western blot, immunoprecipitation, radioimmunoassay, *etc.*

mRNA detection. The diagnostic methods of the invention can also or alternatively involve detection of mRNA encoded by a gene corresponding to a differentially expressed polynucleotides of the invention. Any suitable qualitative or quantitative methods known in the art for detecting specific mRNAs can be used. mRNA can be detected by, for example, *in situ* hybridization in tissue sections, by reverse transcriptase-PCR, or in Northern blots containing poly A⁺ mRNA. One

of skill in the art can readily use these methods to determine differences in the size or amount of mRNA transcripts between two samples. mRNA expression levels in a sample can also be determined by generation of a library of expressed sequence tags (ESTs) from the sample, where the EST library is representative of sequences present in the sample (Adams, et al., (1991) *Science* 252:1651). Enumeration of the relative representation of ESTs within the library can be used to approximate the relative representation of the gene transcript within the starting sample. The results of EST analysis of a test sample can then be compared to EST analysis of a reference sample to determine the relative expression levels of a selected polynucleotide, particularly a polynucleotide corresponding to one or more of the differentially expressed genes described herein. Alternatively, gene expression in a test sample can be performed using serial analysis of gene expression (SAGE) methodology (e.g., Velculescu et al., *Science* (1995) 270:484) or differential display (DD) methodology (see, e.g., U.S. 5,776,683; and U.S. 5,807,680).

Alternatively, gene expression can be analyzed using hybridization analysis. Oligonucleotides or cDNA can be used to selectively identify or capture DNA or RNA of specific sequence composition, and the amount of RNA or cDNA hybridized to a known capture sequence determined qualitatively or quantitatively, to provide information about the relative representation of a particular message within the pool of cellular messages in a sample. Hybridization analysis can be designed to allow for concurrent screening of the relative expression of hundreds to thousands of genes by using, for example, array-based technologies having high density formats, including filters, microscope slides, or microchips, or solution-based technologies that use spectroscopic analysis (e.g., mass spectrometry). One exemplary use of arrays in the diagnostic methods of the invention is described below in more detail.

Use of a single gene in diagnostic applications. The diagnostic methods of the invention can focus on the expression of a single differentially expressed gene. For example, the diagnostic method can involve detecting a differentially expressed gene, or a polymorphism of such a gene (e.g., a polymorphism in an coding region or control region), that is associated with disease. Disease-associated polymorphisms can include deletion or truncation of the gene, mutations that alter expression level and/or affect activity of the encoded protein, etc.

A number of methods are available for analyzing nucleic acids for the presence of a specific sequence, e.g. a disease associated polymorphism. Where large amounts of DNA are available, genomic DNA is used directly. Alternatively, the region of interest is cloned into a suitable vector and grown in sufficient quantity for analysis. Cells that express a differentially expressed gene can be used as a source of mRNA, which can be assayed directly or reverse transcribed into cDNA for analysis. The nucleic acid can be amplified by conventional techniques, such as the polymerase chain reaction (PCR), to provide sufficient amounts for analysis, and a detectable label can be included in the amplification reaction (e.g., using a detectably labeled

primer or detectably labeled oligonucleotides) to facilitate detection. Alternatively, various methods are also known in the art that utilize oligonucleotide ligation as a means of detecting polymorphisms, see e.g., Riley *et al.*, *Nucl. Acids Res.* (1990) 18:2887; and Delahunty *et al.*, *Am. J. Hum. Genet.* (1996) 58:1239.

5 The amplified or cloned sample nucleic acid can be analyzed by one of a number of methods known in the art. The nucleic acid can be sequenced by dideoxy or other methods, and the sequence of bases compared to a selected sequence, e.g., to a wild-type sequence. Hybridization with the polymorphic or variant sequence can also be used to determine its presence in a sample (e.g., by Southern blot, dot blot, *etc.*). The hybridization pattern of a polymorphic or variant
10 sequence and a control sequence to an array of oligonucleotide probes immobilized on a solid support, as described in US 5,445,934, or in WO 95/35505, can also be used as a means of identifying polymorphic or variant sequences associated with disease. Single strand conformational polymorphism (SSCP) analysis, denaturing gradient gel electrophoresis (DGGE), and heteroduplex analysis in gel matrices are used to detect conformational changes created by DNA sequence
15 variation as alterations in electrophoretic mobility. Alternatively, where a polymorphism creates or destroys a recognition site for a restriction endonuclease, the sample is digested with that endonuclease, and the products size fractionated to determine whether the fragment was digested. Fractionation is performed by gel or capillary electrophoresis, particularly acrylamide or agarose gels.

20 Screening for mutations in a gene can be based on the functional or antigenic characteristics of the protein. Protein truncation assays are useful in detecting deletions that can affect the biological activity of the protein. Various immunoassays designed to detect polymorphisms in proteins can be used in screening. Where many diverse genetic mutations lead to a particular disease phenotype, functional protein assays have proven to be effective screening
25 tools. The activity of the encoded protein can be determined by comparison with the wild-type protein.

Pattern matching in diagnosis using arrays. In another embodiment, the diagnostic and/or prognostic methods of the invention involve detection of expression of a selected set of genes in a test sample to produce a test expression pattern (TEP). The TEP is compared to a reference
30 expression pattern (REP), which is generated by detection of expression of the selected set of genes in a reference sample (e.g., a positive or negative control sample). The selected set of genes includes at least one of the genes of the invention, which genes correspond to the polynucleotide sequences of SEQ ID NOS:1-316. Of particular interest is a selected set of genes that includes gene differentially expressed in the disease for which the test sample is to be screened.

35 "Reference sequences" or "reference polynucleotides" as used herein in the context of differential gene expression analysis and diagnosis/prognosis refers to a selected set of

polynucleotides, which selected set includes at least one or more of the differentially expressed polynucleotides described herein. A plurality of reference sequences, preferably comprising positive and negative control sequences, can be included as reference sequences. Additional suitable reference sequences are found in GenBank, Unigene, and other nucleotide sequence databases (including, *e.g.*, expressed sequence tag (EST), partial, and full-length sequences).

"Reference array" means an array having reference sequences for use in hybridization with a sample, where the reference sequences include all, at least one of, or any subset of the differentially expressed polynucleotides described herein. Usually such an array will include at least 3 different reference sequences, and can include any one or all of the provided differentially expressed sequences. Arrays of interest can further comprise sequences, including polymorphisms, of other genetic sequences, particularly other sequences of interest for screening for a disease or disorder (*e.g.*, cancer, dysplasia, or other related or unrelated diseases, disorders, or conditions). The oligonucleotide sequence on the array will usually be at least about 12 nt in length, and can be of about the length of the provided sequences, or can extend into the flanking regions to generate fragments of 100 nt to 200 nt in length or more. Reference arrays can be produced according to any suitable methods known in the art. For example, methods of producing large arrays of oligonucleotides are described in U.S. 5,134,854, and U.S. 5,445,934 using light-directed synthesis techniques. Using a computer controlled system, a heterogeneous array of monomers is converted, through simultaneous coupling at a number of reaction sites, into a heterogeneous array of polymers. Alternatively, microarrays are generated by deposition of pre-synthesized oligonucleotides onto a solid substrate, for example as described in PCT published application no. WO 95/35505.

A "reference expression pattern" or "REP" as used herein refers to the relative levels of expression of a selected set of genes, particularly of differentially expressed genes, that is associated with a selected cell type, *e.g.*, a normal cell, a cancerous cell, a cell exposed to an environmental stimulus, and the like. A "test expression pattern" or "TEP" refers to relative levels of expression of a selected set of genes, particularly of differentially expressed genes, in a test sample (*e.g.*, a cell of unknown or suspected disease state, from which mRNA is isolated).

REPs can be generated in a variety of ways according to methods well known in the art. For example, REPs can be generated by hybridizing a control sample to an array having a selected set of polynucleotides (particularly a selected set of differentially expressed polynucleotides), acquiring the hybridization data from the array, and storing the data in a format that allows for ready comparison of the REP with a TEP. Alternatively, all expressed sequences in a control sample can be isolated and sequenced, *e.g.*, by isolating mRNA from a control sample, converting the mRNA into cDNA, and sequencing the cDNA. The resulting sequence information roughly or precisely reflects the identity and relative number of expressed sequences in the sample. The

sequence information can then be stored in a format (*e.g.*, a computer-readable format) that allows for ready comparison of the REP with a TEP. The REP can be normalized prior to or after data storage, and/or can be processed to selectively remove sequences of expressed genes that are of less interest or that might complicate analysis (*e.g.*, some or all of the sequences associated with housekeeping genes can be eliminated from REP data).

TEPs can be generated in a manner similar to REPs, *e.g.*, by hybridizing a test sample to an array having a selected set of polynucleotides, particularly a selected set of differentially expressed polynucleotides, acquiring the hybridization data from the array, and storing the data in a format that allows for ready comparison of the TEP with a REP. The REP and TEP to be used in a comparison can be generated simultaneously, or the TEP can be compared to previously generated and stored REPs.

In one embodiment of the invention, comparison of a TEP with a REP involves hybridizing a test sample with a reference array, where the reference array has one or more reference sequences for use in hybridization with a sample. The reference sequences include all, at least one of, or any subset of the differentially expressed polynucleotides described herein. Hybridization data for the test sample is acquired, the data normalized, and the produced TEP compared with a REP generated using an array having the same or similar selected set of differentially expressed polynucleotides. Probes that correspond to sequences differentially expressed between the two samples will show decreased or increased hybridization efficiency for one of the samples relative to the other.

Methods for collection of data from hybridization of samples with a reference arrays are well known in the art. For example, the polynucleotides of the reference and test samples can be generated using a detectable fluorescent label, and hybridization of the polynucleotides in the samples detected by scanning the microarrays for the presence of the detectable label using, for example, a microscope and light source for directing light at a substrate. A photon counter detects fluorescence from the substrate, while an x-y translation stage varies the location of the substrate. A confocal detection device that can be used in the subject methods is described in USPN 5,631,734. A scanning laser microscope is described in Shalon et al., *Genome Res.* (1996) 6:639. A scan, using the appropriate excitation line, is performed for each fluorophore used. The digital images generated from the scan are then combined for subsequent analysis. For any particular array element, the ratio of the fluorescent signal from one sample (*e.g.*, a test sample) is compared to the fluorescent signal from another sample (*e.g.*, a reference sample), and the relative signal intensity determined.

Methods for analyzing the data collected from hybridization to arrays are well known in the art. For example, where detection of hybridization involves a fluorescent label, data analysis can include the steps of determining fluorescent intensity as a function of substrate position from the data collected, removing outliers, *i.e.* data deviating from a predetermined statistical distribution,

and calculating the relative binding affinity of the targets from the remaining data. The resulting data can be displayed as an image with the intensity in each region varying according to the binding affinity between targets and probes.

In general, the test sample is classified as having a gene expression profile corresponding to that associated with a disease or non-disease state by comparing the TEP generated from the test sample to one or more REPs generated from reference samples (*e.g.*, from samples associated with cancer or specific stages of cancer, dysplasia, samples affected by a disease other than cancer, normal samples, *etc.*). The criteria for a match or a substantial match between a TEP and a REP include expression of the same or substantially the same set of reference genes, as well as expression of these reference genes at substantially the same levels (*e.g.*, no significant difference between the samples for a signal associated with a selected reference sequence after normalization of the samples, or at least no greater than about 25% to about 40% difference in signal strength for a given reference sequence. In general, a pattern match between a TEP and a REP includes a match in expression, preferably a match in qualitative or quantitative expression level, of at least one of, all or any subset of the differentially expressed genes of the invention.

Pattern matching can be performed manually, or can be performed using a computer program. Methods for preparation of substrate matrices (*e.g.*, arrays), design of oligonucleotides for use with such matrices, labeling of probes, hybridization conditions, scanning of hybridized matrices, and analysis of patterns generated, including comparison analysis, are described in, for example, U.S. 5,800,992.

Diagnosis, Prognosis and Management of Cancer

The polynucleotides of the invention and their gene products are of particular interest as genetic or biochemical markers (*e.g.*, in blood or tissues) that will detect the earliest changes along the carcinogenesis pathway and/or to monitor the efficacy of various therapies and preventive interventions. For example, the level of expression of certain polynucleotides can be indicative of a poorer prognosis, and therefore warrant more aggressive chemo- or radio-therapy for a patient or vice versa. The correlation of novel surrogate tumor specific features with response to treatment and outcome in patients can define prognostic indicators that allow the design of tailored therapy based on the molecular profile of the tumor. These therapies include antibody targeting and gene therapy. Determining expression of certain polynucleotides and comparison of a patient's profile with known expression in normal tissue and variants of the disease allows a determination of the best possible treatment for a patient, both in terms of specificity of treatment and in terms of comfort level of the patient. Surrogate tumor markers, such as polynucleotide expression, can also be used to better classify, and thus diagnose and treat, different forms and disease states of cancer. Two classifications widely used in oncology that can benefit from identification of the expression

levels of the polynucleotides of the invention are staging of the cancerous disorder, and grading the nature of the cancerous tissue.

The polynucleotides of the invention can be useful to monitor patients having or susceptible to cancer to detect potentially malignant events at a molecular level before they are detectable at a gross morphological level. Furthermore, a polynucleotide of the invention identified as important for one type of cancer can also have implications for development or risk of development of other types of cancer, e.g., where a polynucleotide is differentially expressed across various cancer types. Thus, for example, expression of a polynucleotide that has clinical implications for metastatic colon cancer can also have clinical implications for stomach cancer or endometrial cancer.

Staging. Staging is a process used by physicians to describe how advanced the cancerous state is in a patient. Staging assists the physician in determining a prognosis, planning treatment and evaluating the results of such treatment. Staging systems vary with the types of cancer, but generally involve the following “TNM” system: the type of tumor, indicated by T; whether the cancer has metastasized to nearby lymph nodes, indicated by N; and whether the cancer has metastasized to more distant parts of the body, indicated by M. Generally, if a cancer is only detectable in the area of the primary lesion without having spread to any lymph nodes it is called Stage I. If it has spread only to the closest lymph nodes, it is called Stage II. In Stage III, the cancer has generally spread to the lymph nodes in near proximity to the site of the primary lesion. Cancers that have spread to a distant part of the body, such as the liver, bone, brain or other site, are Stage IV, the most advanced stage.

The polynucleotides of the invention can facilitate fine-tuning of the staging process by identifying markers for the aggressivity of a cancer, e.g. the metastatic potential, as well as the presence in different areas of the body. Thus, a Stage II cancer with a polynucleotide signifying a high metastatic potential cancer can be used to change a borderline Stage II tumor to a Stage III tumor, justifying more aggressive therapy. Conversely, the presence of a polynucleotide signifying a lower metastatic potential allows more conservative staging of a tumor.

Grading of cancers. Grade is a term used to describe how closely a tumor resembles normal tissue of its same type. The microscopic appearance of a tumor is used to identify tumor grade based on parameters such as cell morphology, cellular organization, and other markers of differentiation. As a general rule, the grade of a tumor corresponds to its rate of growth or aggressiveness, with undifferentiated or high-grade tumors being more aggressive than well differentiated or low-grade tumors. The following guidelines are generally used for grading tumors: 1) GX Grade cannot be assessed; 2) G1 Well differentiated; G2 Moderately well differentiated; 3) G3 Poorly differentiated; 4) G4 Undifferentiated. The polynucleotides of the invention can be especially valuable in determining the grade of the tumor, as they not only can aid in determining

the differentiation status of the cells of a tumor, they can also identify factors other than differentiation that are valuable in determining the aggressiveness of a tumor, such as metastatic potential.

Detection of lung cancer. The polynucleotides of the invention can be used to detect lung cancer in a subject. Although there are more than a dozen different kinds of lung cancer, the two main types of lung cancer are small cell and nonsmall cell, which encompass about 90% of all lung cancer cases. Small cell carcinoma (also called oat cell carcinoma) usually starts in one of the larger bronchial tubes, grows fairly rapidly, and is likely to be large by the time of diagnosis. Nonsmall cell lung cancer (NSCLC) is made up of three general subtypes of lung cancer. Epidermoid carcinoma (also called squamous cell carcinoma) usually starts in one of the larger bronchial tubes and grows relatively slowly. The size of these tumors can range from very small to quite large. Adenocarcinoma starts growing near the outside surface of the lung and can vary in both size and growth rate. Some slowly growing adenocarcinomas are described as alveolar cell cancer. Large cell carcinoma starts near the surface of the lung, grows rapidly, and the growth is usually fairly large when diagnosed. Other less common forms of lung cancer are carcinoid, cylindroma, mucoepidermoid, and malignant mesothelioma.

The polynucleotides of the invention, e.g., polynucleotides differentially expressed in normal cells versus cancerous lung cells (e.g., tumor cells of high or low metastatic potential) or between types of cancerous lung cells (e.g., high metastatic versus low metastatic), can be used to distinguish types of lung cancer as well as identifying traits specific to a certain patient's cancer and selecting an appropriate therapy. For example, if the patient's biopsy expresses a polynucleotide that is associated with a low metastatic potential, it may justify leaving a larger portion of the patient's lung in surgery to remove the lesion. Alternatively, a smaller lesion with expression of a polynucleotide that is associated with high metastatic potential may justify a more radical removal of lung tissue and/or the surrounding lymph nodes, even if no metastasis can be identified through pathological examination.

Detection of breast cancer. The majority of breast cancers are adenocarcinomas subtypes, which can be summarized as follows: 1) ductal carcinoma in situ (DCIS), including comedocarcinoma; 2) infiltrating (or invasive) ductal carcinoma (IDC); 3) lobular carcinoma in situ (LCIS); 4) infiltrating (or invasive) lobular carcinoma (ILC); 5) inflammatory breast cancer; 6) medullary carcinoma; 7) mucinous carcinoma; 8) Paget's disease of the nipple; 9) Phyllodes tumor; and 10) tubular carcinoma;

The expression of polynucleotides of the invention can be used in the diagnosis and management of breast cancer, as well as to distinguish between types of breast cancer. Detection of breast cancer can be determined using expression levels of any of the appropriate polynucleotides of the invention, either alone or in combination. Determination of the aggressive nature and/or the

metastatic potential of a breast cancer can also be determined by comparing levels of one or more polynucleotides of the invention and comparing levels of another sequence known to vary in cancerous tissue, *e.g.* ER expression. In addition, development of breast cancer can be detected by examining the ratio of expression of a differentially expressed polynucleotide to the levels of steroid hormones (*e.g.*, testosterone or estrogen) or to other hormones (*e.g.*, growth hormone, insulin). Thus expression of specific marker polynucleotides can be used to discriminate between normal and cancerous breast tissue, to discriminate between breast cancers with different cells of origin, to discriminate between breast cancers with different potential metastatic rates, etc.

Detection of colon cancer. The polynucleotides of the invention exhibiting the appropriate expression pattern can be used to detect colon cancer in a subject. Colorectal cancer is one of the most common neoplasms in humans and perhaps the most frequent form of hereditary neoplasia. Prevention and early detection are key factors in controlling and curing colorectal cancer. Colorectal cancer begins as polyps, which are small, benign growths of cells that form on the inner lining of the colon. Over a period of several years, some of these polyps accumulate additional mutations and become cancerous. Multiple familial colorectal cancer disorders have been identified, which are summarized as follows: 1) Familial adenomatous polyposis (FAP); 2) Gardner's syndrome; 3) Hereditary nonpolyposis colon cancer (HNPCC); and 4) Familial colorectal cancer in Ashkenazi Jews. The expression of appropriate polynucleotides of the invention can be used in the diagnosis, prognosis and management of colorectal cancer. Detection of colon cancer can be determined using expression levels of any of these sequences alone or in combination with the levels of expression. Determination of the aggressive nature and/or the metastatic potential of a colon cancer can be determined by comparing levels of one or more polynucleotides of the invention and comparing total levels of another sequence known to vary in cancerous tissue, *e.g.*, expression of p53, DCC ras, or FAP (see, *e.g.*, Fearon ER, *et al.*, *Cell* (1990) 61(5):759; Hamilton SR *et al.*, *Cancer* (1993) 72:957; Bodmer W, *et al.*, *Nat Genet.* (1994) 4(3):217; Fearon ER, *Ann N Y Acad Sci.* (1995) 768:101). For example, development of colon cancer can be detected by examining the ratio of any of the polynucleotides of the invention to the levels of oncogenes (*e.g.* ras) or tumor suppressor genes (*e.g.* FAP or p53). Thus expression of specific marker polynucleotides can be used to discriminate between normal and cancerous colon tissue, to discriminate between colon cancers with different cells of origin, to discriminate between colon cancers with different potential metastatic rates, etc.

Detection of prostate cancer. The polynucleotides and their corresponding genes and gene products exhibiting the appropriate differential expression pattern can be used to detect prostate cancer in a subject. Over 95% of primary prostate cancers are adenocarcinomas. Signs and symptoms may include: frequent urination, especially at night, inability to urinate, trouble starting

or holding back urination, a weak or interrupted urine flow and frequent pain or stiffness in the lower back, hips or upper thighs.

Many of the signs and symptoms of prostate cancer can be caused by a variety of other non-cancerous conditions. For example, one common cause of many of these signs and symptoms is a condition called benign prostatic hypertrophy, or BPH. In BPH, the prostate gets bigger and may block the flow of urine or interfere with sexual function. The methods and compositions of the invention can be used to distinguish between prostate cancer and such non-cancerous conditions. The methods of the invention can be used in conjunction with conventional methods of diagnosis, e.g., digital rectal exam and/or detection of the level of prostate specific antigen (PSA), a substance produced and secreted by the prostate.

Use of Polynucleotides to Screen for Peptide Analogs and Antagonists

Polypeptides encoded by the instant polynucleotides and corresponding full length genes can be used to screen peptide libraries to identify binding partners, such as receptors, from among the encoded polypeptides. Peptide libraries can be synthesized according to methods known in the art (see, e.g., USPN 5,010,175, and WO 91/17823). Agonists or antagonists of the polypeptides if the invention can be screened using any available method known in the art, such as signal transduction, antibody binding, receptor binding, mitogenic assays, chemotaxis assays, etc. The assay conditions ideally should resemble the conditions under which the native activity is exhibited *in vivo*, that is, under physiologic pH, temperature, and ionic strength. Suitable agonists or antagonists will exhibit strong inhibition or enhancement of the native activity at concentrations that do not cause toxic side effects in the subject. Agonists or antagonists that compete for binding to the native polypeptide can require concentrations equal to or greater than the native concentration, while inhibitors capable of binding irreversibly to the polypeptide can be added in concentrations on the order of the native concentration.

Such screening and experimentation can lead to identification of a novel polypeptide binding partner, such as a receptor, encoded by a gene or a cDNA corresponding to a polynucleotide of the invention, and at least one peptide agonist or antagonist of the novel binding partner. Such agonists and antagonists can be used to modulate, enhance, or inhibit receptor function in cells to which the receptor is native, or in cells that possess the receptor as a result of genetic engineering. Further, if the novel receptor shares biologically important characteristics with a known receptor, information about agonist/antagonist binding can facilitate development of improved agonists/antagonists of the known receptor.

Pharmaceutical Compositions and Therapeutic Uses

Pharmaceutical compositions of the invention can comprise polypeptides, antibodies, or polynucleotides (including antisense nucleotides and ribozymes) of the claimed invention in a therapeutically effective amount. The term "therapeutically effective amount" as used herein refers

to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature. The precise effective amount for a subject will
5 depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation is determined by routine experimentation and is within the judgment of the clinician. For purposes of the present invention, an effective dose will generally be from about 0.01 mg/ kg to 50 mg/kg or
10 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the
15 individual receiving the composition, and which can be administered without undue toxicity. Suitable carriers can be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Pharmaceutically acceptable carriers in therapeutic compositions can include liquids such
20 as water, saline, glycerol and ethanol. Auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, can also be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.
25 Pharmaceutically acceptable salts can also be present in the pharmaceutical composition, e.g., mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in *Remington's Pharmaceutical Sciences* (Mack Pub. Co., N.J. 1991).

30 Delivery Methods. Once formulated, the compositions of the invention can be (1) administered directly to the subject (e.g., as polynucleotide or polypeptides); or (2) delivered ex vivo, to cells derived from the subject (e.g., as in *ex vivo* gene therapy). Direct delivery of the compositions will generally be accomplished by parenteral injection, e.g., subcutaneously, intraperitoneally, intravenously or intramuscularly, intratumoral or to the interstitial space of a
35 tissue. Other modes of administration include oral and pulmonary administration, suppositories,

and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment can be a single dose schedule or a multiple dose schedule.

Methods for the ex vivo delivery and reimplantation of transformed cells into a subject are known in the art and described in *e.g.*, International Publication No. WO 93/14778. Examples of cells useful in ex vivo applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells. Generally, delivery of nucleic acids for both ex vivo and in vitro applications can be accomplished by, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Once a gene corresponding to a polynucleotide of the invention has been found to correlate with a proliferative disorder, such as neoplasia, dysplasia, and hyperplasia, the disorder can be amenable to treatment by administration of a therapeutic agent based on the provided polynucleotide, corresponding polypeptide or other corresponding molecule (*e.g.*, antisense, ribozyme, etc.).

The dose and the means of administration of the inventive pharmaceutical compositions are determined based on the specific qualities of the therapeutic composition, the condition, age, and weight of the patient, the progression of the disease, and other relevant factors. For example, administration of polynucleotide therapeutic compositions agents of the invention includes local or systemic administration, including injection, oral administration, particle gun or catheterized administration, and topical administration. Preferably, the therapeutic polynucleotide composition contains an expression construct comprising a promoter operably linked to a polynucleotide of at least 12, 22, 25, 30, or 35 contiguous nt of the polynucleotide disclosed herein. Various methods can be used to administer the therapeutic composition directly to a specific site in the body. For example, a small metastatic lesion is located and the therapeutic composition injected several times in several different locations within the body of tumor. Alternatively, arteries which serve a tumor are identified, and the therapeutic composition injected into such an artery, in order to deliver the composition directly into the tumor. A tumor that has a necrotic center is aspirated and the composition injected directly into the now empty center of the tumor. The antisense composition is directly administered to the surface of the tumor, for example, by topical application of the composition. X-ray imaging is used to assist in certain of the above delivery methods.

Receptor-mediated targeted delivery of therapeutic compositions containing an antisense polynucleotide, subgenomic polynucleotides, or antibodies to specific tissues can also be used. Receptor-mediated DNA delivery techniques are described in, for example, Findeis *et al.*, *Trends Biotechnol.* (1993) 11:202; Chiou *et al.*, *Gene Therapeutics: Methods And Applications Of Direct Gene Transfer* (J.A. Wolff, ed.) (1994); Wu *et al.*, *J. Biol. Chem.* (1988) 263:621; Wu *et al.*, *J. Biol.*

Chem. (1994) 269:542; Zenke *et al.*, *Proc. Natl. Acad. Sci. (USA)* (1990) 87:3655; Wu *et al.*, *J. Biol. Chem.* (1991) 266:338. Therapeutic compositions containing a polynucleotide are administered in a range of about 100 ng to about 200 mg of DNA for local administration in a gene therapy protocol. Concentration ranges of about 500 ng to about 50 mg, about 1 µg to about 2 mg, about 5 µg to about 500 µg, and about 20 µg to about 100 µg of DNA can also be used during a gene therapy protocol. Factors such as method of action (e.g., for enhancing or inhibiting levels of the encoded gene product) and efficacy of transformation and expression are considerations which will affect the dosage required for ultimate efficacy of the antisense subgenomic polynucleotides. Where greater expression is desired over a larger area of tissue, larger amounts of antisense subgenomic polynucleotides or the same amounts readministered in a successive protocol of administrations, or several administrations to different adjacent or close tissue portions of, for example, a tumor site, may be required to effect a positive therapeutic outcome. In all cases, routine experimentation in clinical trials will determine specific ranges for optimal therapeutic effect. For polynucleotide related genes encoding polypeptides or proteins with anti-inflammatory activity, suitable use, doses, and administration are described in USPN 5,654,173.

The therapeutic polynucleotides and polypeptides of the present invention can be delivered using gene delivery vehicles. The gene delivery vehicle can be of viral or non-viral origin (see generally, Jolly, *Cancer Gene Therapy* (1994) 1:51; Kimura, *Human Gene Therapy* (1994) 5:845; Connelly, *Human Gene Therapy* (1995) 1:185; and Kaplitt, *Nature Genetics* (1994) 6:148).

Expression of such coding sequences can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence can be either constitutive or regulated.

Viral-based vectors for delivery of a desired polynucleotide and expression in a desired cell are well known in the art. Exemplary viral-based vehicles include, but are not limited to, recombinant retroviruses (see, e.g., WO 90/07936; WO 94/03622; WO 93/25698; WO 93/25234; USPN 5, 219,740; WO 93/11230; WO 93/10218; USPN 4,777,127; GB Patent No. 2,200,651; EP 0 345 242; and WO 91/02805), alphavirus-based vectors (e.g., Sindbis virus vectors, Semliki forest virus (ATCC VR-67; ATCC VR-1247), Ross River virus (ATCC VR-373; ATCC VR-1246) and Venezuelan equine encephalitis virus (ATCC VR-923; ATCC VR-1250; ATCC VR 1249; ATCC VR-532), and adeno-associated virus (AAV) vectors (see, e.g., WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 and WO 95/00655). Administration of DNA linked to killed adenovirus as described in Curiel, *Hum. Gene Ther.* (1992) 3:147 can also be employed.

Non-viral delivery vehicles and methods can also be employed, including, but not limited to, polycationic condensed DNA linked or unlinked to killed adenovirus alone (see, e.g., Curiel, *Hum. Gene Ther.* (1992) 3:147); ligand-linked DNA (see, e.g., Wu, *J. Biol. Chem.* (1989) 264:16985); eukaryotic cell delivery vehicles (see, e.g., USPN 5,814,482; WO 95/07994; WO 96/17072; WO 95/30763; and WO 97/42338) and nucleic charge neutralization or fusion with

cell membranes. Naked DNA can also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and USPN 5,580,859. Liposomes that can act as gene delivery vehicles are described in USPN 5,422,120; WO 95/13796; WO 94/23697; WO 91/14445; and EP 0524968. Additional approaches are described in Philip, *Mol. Cell Biol.* (1994) 14:2411, and in
5 Woffendin, *Proc. Natl. Acad. Sci.* (1994) 91:1581

Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al.*, *Proc. Natl. Acad. Sci. USA* (1994) 91(24):11581. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials or use of ionizing radiation (see, e.g., USPN
10 5,206,152 and WO 92/11033). Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun (see, e.g., USPN 5,149,655); use of ionizing radiation for activating transferred gene (see, e.g., USPN 5,206,152 and WO 92/11033).

The present invention will now be illustrated by reference to the following examples which
15 set forth particularly advantageous embodiments. However, it should be noted that these embodiments are illustrative and are not to be construed as restricting the invention in any way.

EXAMPLES

The following examples are offered primarily for purposes of illustration. It will be readily
20 apparent to those skilled in the art that the formulations, dosages, methods of administration, and other parameters of this invention may be further modified or substituted in various ways without departing from the spirit and scope of the invention.

Example 1: Source of Biological Materials and Overview of Novel Polynucleotides Expressed 25 by the Biological Materials

cDNA libraries were constructed from mRNA isolated from the GRRpz or and WOca cells, which were provided by Dr. Donna M. Peehl, Department of Medicine, Stanford University School of Medicine. GRRpz cells were primary cells derived from normal prostate epithelium. The WOca cells were prostate epithelial cells derived from prostate cancer Gleason Grade 4+4.
30 Polynucleotides expressed by these cells were isolated and analyzed; the sequences of these polynucleotides were about 275-300 nucleotides in length.

The sequences of the isolated polynucleotides were first masked to eliminate low complexity sequences using the XBLAST masking program (Claverie "Effective Large-Scale Sequence Similarity Searches," In: Computer Methods for Macromolecular Sequence Analysis,
35 Doolittle, ed., *Meth. Enzymol.* 266:212-227 Academic Press, NY, NY (1996); see particularly Claverie, in "Automated DNA Sequencing and Analysis Techniques" Adams *et al.*, eds., Chap. 36,

p. 267 Academic Press, San Diego, 1994 and Claverie *et al. Comput. Chem.* (1993) 17:191).

Generally, masking does not influence the final search results, except to eliminate sequences of relative little interest due to their low complexity, and to eliminate multiple "hits" based on similarity to repetitive regions common to multiple sequences, e.g., Alu repeats. The remaining
5 sequences were then used in a BLASTN vs. GenBank search; sequences that exhibited greater than 70% overlap, 99% identity, and a p value of less than 1×10^{-40} were discarded. Sequences from this search also were discarded if the inclusive parameters were met, but the sequence was ribosomal or vector-derived.

The resulting sequences from the previous search were classified into three groups (1, 2
10 and 3 below) and searched in a BLASTX vs. NRP (non-redundant proteins) database search: (1) unknown (no hits in the GenBank search), (2) weak similarity (greater than 45% identity and p value of less than 1×10^{-5}), and (3) high similarity (greater than 60% overlap, greater than 80% identity, and p value less than 1×10^{-5}). Sequences having greater than 70% overlap, greater than 99% identity, and p value of less than 1×10^{-40} were discarded.

15 The remaining sequences were classified as unknown (no hits), weak similarity, and high similarity (parameters as above). Two searches were performed on these sequences. First, a BLAST vs. EST database search was performed and sequences with greater than 99% overlap, greater than 99% similarity and a p value of less than 1×10^{-40} were discarded. Sequences with a p value of less than 1×10^{-65} when compared to a database sequence of human origin were also
20 excluded. Second, a BLASTN vs. Patent GeneSeq database was performed and sequences having greater than 99% identity, p value less than 1×10^{-40} , and greater than 99% overlap were discarded.

The remaining sequences were subjected to screening using other rules and redundancies in the dataset. Sequences with a p value of less than 1×10^{-111} in relation to a database sequence of human origin were specifically excluded. The final result provided the 316 sequences listed as
25 SEQ ID NOS:1-316 in the accompanying Sequence Listing and summarized in Table 1 (inserted prior to claims). Each identified polynucleotide represents sequence from at least a partial mRNA transcript. Many of the sequences include the sequence ggcacgag at the 5' end; this sequence is a sequencing artifact and not part of the sequence of the polynucleotides of the invention.

Table 1 provides: 1) the SEQ ID NO ("SEQ ID") assigned to each sequence for use in the
30 present specification; 2) the Cluster Identification No. ("CLUSTER"); 3) the sequence name ("SEQ NAME") used as an internal identifier of the sequence; 4) the orientation of the sequence ("ORIENT"); 5) the name assigned to the clone from which the sequence was isolated ("CLONE ID"); and the name of the library from which the sequence was isolated ("LIBRARY"). CH22PRC indicates the sequence was isolated from Library 22; CH21PRN indicates the sequence was isolated
35 from Library 21. A description of the libraries is provided in Table 3 below. Because the provided polynucleotides represent partial mRNA transcripts, two or more polynucleotides of the invention

may represent different regions of the same mRNA transcript and the same gene. Thus, if two or more SEQ ID NOS: are identified as belonging to the same clone, then either sequence can be used to obtain the full-length mRNA or gene.

5 Example 2: Results of Public Database Search to Identify Function of Gene Products

SEQ ID NOS:1-316 were translated in all three reading frames, and the nucleotide sequences and translated amino acid sequences used as query sequences to search for homologous sequences in either the GenBank (nucleotide sequences) or Non-Redundant Protein (amino acid sequences) databases. Query and individual sequences were aligned using the BLAST 2.0
10 programs, available over the world wide web at a saite sponsored by the National Center for Biotechnology Information, which is supported by the National Library of Medicine and the National Institutes of Health (see also Altschul, et al. *Nucleic Acids Res.* (1997) 25:3389-3402). The sequences were masked to various extents to prevent searching of repetitive sequences or poly-A sequences, using the XBLAST program for masking low complexity as described above in
15 Example 1.

Table 2 (inserted before the claims) provide the alignment summaries having a p value of 1×10^{-2} or less indicating substantial homology between the sequences of the present invention and those of the indicated public databases. Specifically, Table 2 provides the SEQ ID NO of the query sequence, the accession number of the GenBank database entry of the homologous sequence, and
20 the p value of the alignment. Table 2 also provides the SEQ ID NO of the query sequence, the accession number of the Non-Redundant Protein database entry of the homologous sequence, and the p value of the alignment. The alignments provided in Table 2 are the best available alignment to a DNA or amino acid sequence at a time just prior to filing of the present specification. The activity of the polypeptide encoded by the SEQ ID NOS listed in Table 2 can be extrapolated to be
25 substantially the same or substantially similar to the activity of the reported nearest neighbor or closely related sequence. The accession number of the nearest neighbor is reported, providing a publicly available reference to the activities and functions exhibited by the nearest neighbor. The public information regarding the activities and functions of each of the nearest neighbor sequences is incorporated by reference in this application. Also incorporated by reference is all publicly
30 available information regarding the sequence, as well as the putative and actual activities and functions of the nearest neighbor sequences listed in Table 2 and their related sequences. The search program and database used for the alignment, as well as the calculation of the p value are also indicated.

Full length sequences or fragments of the polynucleotide sequences of the nearest
35 neighbors can be used as probes and primers to identify and isolate the full length sequence of the corresponding polynucleotide. The nearest neighbors can indicate a tissue or cell type to be used to

construct a library for the full-length sequences of the corresponding polynucleotides.

Example 3: Differential Expression of Polynucleotides of the Invention: Description of Libraries and

5 Detection of Differential Expression

The relative expression levels of the polynucleotides of the invention was assessed in several libraries prepared from various sources, including primary cells, cell lines and patient tissue samples. Table 3 provides a summary of these libraries, including the shortened library name (used hereafter), the mRNA source used to prepared the cDNA library, the "nickname" of the library that
10 is used in the tables below (in quotes), and the approximate number of clones in the library.

Table 3. Description of cDNA Libraries

Library (Lib#)	Description	Number of Clones in Library
1	Human Colon Cell Line Km12 L4: High Metastatic Potential (derived from Km12C)	308731
2	Human Colon Cell Line Km12C: Low Metastatic Potential	284771
3	Human Breast Cancer Cell Line MDA-MB-231: High Metastatic Potential; micro-mets in lung	326937
4	Human Breast Cancer Cell Line MCF7: Non Metastatic	318979
8	Human Lung Cancer Cell Line MV-522: High Metastatic Potential	223620
9	Human Lung Cancer Cell Line UCP-3: Low Metastatic Potential	312503
12	Human microvascular endothelial cells (HMVEC) - UNTREATED (PCR (OligodT) cDNA library)	41938
13	Human microvascular endothelial cells (HMVEC) – bFGF TREATED (PCR (OligodT) cDNA library)	42100
14	Human microvascular endothelial cells (HMVEC) – VEGF TREATED (PCR (OligodT) cDNA library)	42825
15	Normal Colon - UC#2 Patient (MICRODISSECTED PCR (OligodT) cDNA library)	282722
16	Colon Tumor - UC#2 Patient (MICRODISSECTED PCR (OligodT) cDNA library)	298831
17	Liver Metastasis from Colon Tumor of UC#2 Patient	303467

Library (Lib#)	Description	Number of Clones in Library
	(MICRODISSECTED PCR (OligodT) cDNA library)	
18	Normal Colon - UC#3 Patient (MICRODISSECTED PCR (OligodT) cDNA library)	36216
19	Colon Tumor - UC#3 Patient (MICRODISSECTED PCR (OligodT) cDNA library)	41388
20	Liver Metastasis from Colon Tumor of UC#3 Patient (MICRODISSECTED PCR (OligodT) cDNA library)	30956
21	GRRpz Cells derived from normal prostate epithelium	164801
22	WOca Cells derived from Gleason Grade 4 prostate cancer epithelium	162088
23-	Normal Lung Epithelium of Patient #1006 (MICRODISSECTED PCR (OligodT) cDNA library)	306198
24	Primary tumor, Large Cell Carcinoma of Patient #1006 (MICRODISSECTED PCR (OligodT) cDNA library)	309349

The KM12L4 cell line is derived from the KM12C cell line (Morikawa, et al., *Cancer Research* (1988) 48:6863). The KM12C cell line, which is poorly metastatic (low metastatic) was established in culture from a Dukes' stage B₂ surgical specimen (Morikawa *et al. Cancer Res.* (1988) 48:6863). The KML4-A is a highly metastatic subline derived from KM12C (Yeatman *et al. Nucl. Acids. Res.* (1995) 23:4007; Bao-Ling *et al. Proc. Annu. Meet. Am. Assoc. Cancer. Res.* (1995) 21:3269). The KM12C and KM12C-derived cell lines (*e.g.*, KM12L4, KM12L4-A, *etc.*) are well-recognized in the art as a model cell line for the study of colon cancer (see, *e.g.*, Moriakawa *et al., supra*; Radinsky *et al. Clin. Cancer Res.* (1995) 1:19; Yeatman *et al., (1995) supra*; Yeatman *et al. Clin. Exp. Metastasis* (1996) 14:246). The MDA-MB-231 cell line (Brinkley *et al. Cancer Res.* (1980) 40:3118-3129) was originally isolated from pleural effusions (Cailleau, *J. Natl. Cancer. Inst.* (1974) 53:661), is of high metastatic potential, and forms poorly differentiated adenocarcinoma grade II in nude mice consistent with breast carcinoma.

The MCF7 cell line was derived from a pleural effusion of a breast adenocarcinoma and is non-metastatic. The MV-522 cell line is derived from a human lung carcinoma and is of high metastatic potential. The UCP-3 cell line is a low metastatic human lung carcinoma cell line; the MV-522 is a high metastatic variant of UCP-3. These cell lines are well-recognized in the art as models for the study of human breast and lung cancer (see, *e.g.*, Chandrasekaran *et al., Cancer Res.* (1979) 39:870 (MDA-MB-231 and MCF-7); Gastpar *et al., J Med Chem* (1998) 41:4965 (MDA-

MB-231 and MCF-7); Ranson *et al.*, *Br J Cancer* (1998) 77:1586 (MDA-MB-231 and MCF-7); Kuang *et al.*, *Nucleic Acids Res* (1998) 26:1116 (MDA-MB-231 and MCF-7); Varki *et al.*, *Int J Cancer* (1987) 40:46 (UCP-3); Varki *et al.*, *Tumour Biol.* (1990) 11:327; (MV-522 and UCP-3); Varki *et al.*, *Anticancer Res.* (1990) 10:637; (MV-522); Kelner *et al.*, *Anticancer Res* (1995) 15:867 (MV-522); and Zhang *et al.*, *Anticancer Drugs* (1997) 8:696 (MV522)). The samples of libraries 15-20 are derived from two different patients (UC#2, and UC#3). The bFGF-treated HMVEC were prepared by incubation with bFGF at 10ng/ml for 2 hrs; the VEGF-treated HMVEC were prepared by incubation with 20ng/ml VEGF for 2 hrs. Following incubation with the respective growth factor, the cells were washed and lysis buffer added for RNA preparation. The GRRpz and WOca cells were provided by Dr. Donna M. Peehl, Department of Medicine, Stanford University School of Medicine. GRRpz cells were derived from normal prostate epithelium. The WOca cells are Gleason Grade 4 cell line.

Each of the libraries is composed of a collection of cDNA clones that in turn are representative of the mRNAs expressed in the indicated mRNA source. In order to facilitate the analysis of the millions of sequences in each library, the sequences were assigned to clusters. The concept of "cluster of clones" is derived from a sorting/grouping of cDNA clones based on their hybridization pattern to a panel of roughly 300 7bp oligonucleotide probes (see Drmanac *et al.*, *Genomics* (1996) 37(1):29). Random cDNA clones from a tissue library are hybridized at moderate stringency to 300 7bp oligonucleotides. Each oligonucleotide has some measure of specific hybridization to that specific clone. The combination of 300 of these measures of hybridization for 300 probes equals the "hybridization signature" for a specific clone. Clones with similar sequence will have similar hybridization signatures. By developing a sorting/grouping algorithm to analyze these signatures, groups of clones in a library can be identified and brought together computationally. These groups of clones are termed "clusters". Depending on the stringency of the selection in the algorithm (similar to the stringency of hybridization in a classic library cDNA screening protocol), the "purity" of each cluster can be controlled. For example, artifacts of clustering may occur in computational clustering just as artifacts can occur in "wet-lab" screening of a cDNA library with 400 bp cDNA fragments, at even the highest stringency. The stringency used in the implementation of cluster herein provides groups of clones that are in general from the same cDNA or closely related cDNAs. Closely related clones can be a result of different length clones of the same cDNA, closely related clones from highly related gene families, or splice variants of the same cDNA.

Differential expression for a selected cluster was assessed by first determining the number of cDNA clones corresponding to the selected cluster in the first library (Clones in 1st), and the determining the number of cDNA clones corresponding to the selected cluster in the second library (Clones in 2nd). Differential expression of the selected cluster in the first library relative to the

second library is expressed as a "ratio" of percent expression between the two libraries. In general, the "ratio" is calculated by: 1) calculating the percent expression of the selected cluster in the first library by dividing the number of clones corresponding to a selected cluster in the first library by the total number of clones analyzed from the first library; 2) calculating the percent expression of the selected cluster in the second library by dividing the number of clones corresponding to a selected cluster in a second library by the total number of clones analyzed from the second library; 3) dividing the calculated percent expression from the first library by the calculated percent expression from the second library. If the "number of clones" corresponding to a selected cluster in a library is zero, the value is set at 1 to aid in calculation. The formula used in calculating the ratio takes into account the "depth" of each of the libraries being compared, *i.e.*, the total number of clones analyzed in each library.

In general, a polynucleotide is said to be significantly differentially expressed between two samples when the ratio value is greater than at least about 2, preferably greater than at least about 3, more preferably greater than at least about 5, where the ratio value is calculated using the method described above. The significance of differential expression is determined using a z score test (Zar, Biostatistical Analysis, Prentice Hall, Inc., USA, "Differences between Proportions," pp 296-298 (1974)).

Using this approach, a number of polynucleotide sequences were identified as being differentially expressed between, for example, cells derived from high metastatic potential cancer tissue and low metastatic cancer cells, and between cells derived from metastatic cancer tissue and normal tissue. Evaluation of the levels of expression of the genes corresponding to these sequences can be valuable in diagnosis, prognosis, and/or treatment (*e.g.*, to facilitate rationale design of therapy, monitoring during and after therapy, *etc.*). Moreover, the genes corresponding to differentially expressed sequences described herein can be therapeutic targets due to their involvement in regulation (*e.g.*, inhibition or promotion) of development of, for example, the metastatic phenotype. For example, sequences that correspond to genes that are increased in expression in high metastatic potential cells relative to normal or non-metastatic tumor cells may encode genes or regulatory sequences involved in processes such as angiogenesis, differentiation, cell replication, and metastasis.

Detection of the relative expression levels of differentially expressed polynucleotides described herein can provide valuable information to guide the clinician in the choice of therapy. For example, a patient sample exhibiting an expression level of one or more of these polynucleotides that corresponds to a gene that is increased in expression in metastatic or high metastatic potential cells may warrant more aggressive treatment for the patient. In contrast, detection of expression levels of a polynucleotide sequence that corresponds to expression levels associated with that of low metastatic potential cells may warrant a more positive prognosis than

the gross pathology would suggest.

The differential expression of the polynucleotides described herein can thus be used as, for example, diagnostic markers, prognostic markers, for risk assessment, patient treatment and the like. These polynucleotide sequences can also be used in combination with other known molecular and/or biochemical markers.

The differential expression data for polynucleotides of the invention that have been identified as being differentially expressed across various combinations of the libraries described above is summarized in Table 4 (inserted prior to the claims). Table 4 provides: 1) the Sequence Identification Number ("SEQ ID") assigned to the polynucleotide; 2) the cluster ("CLUST") to which the polynucleotide has been assigned as described above; 3) the library comparisons that resulted in identification of the polynucleotide as being differentially expressed ("PairAB-text"), with shorthand names of the compared libraries provided in parentheses following the library numbers; 4) the number of clones corresponding to the polynucleotide in the first library listed ("A"); 5) the number of clones corresponding to the polynucleotide in the second library listed ("B"); 6) the "RATIO PLUS" where the comparison resulted in a finding that the number of clones in library A is greater than the number of clones in library B; and 7) the "RATIO MINUS" where the comparison resulted in a finding that the number of clones in library B is greater than the number of clones in library A.

Example 4: Differential Expression of a Polynucleotides Associated with Metastatic Potential in Breast Cancer

Differential expression was examined in breast cancer cells having either high metastatic potential or low metastatic potential. A single cluster, Cluster Identification No. 10154, was identified as displaying low expression in the high metastatic potential breast cancer cells (Library 3), and significantly increased expression – approximately 100-fold higher -- in the low metastatic potential cells (Library 4). Specifically, three clones were identified that were expressed in Library 3, the high metastatic potential breast cancer library, while 317 clones were expressed in Library 4, the low metastatic potential breast cancer library. The two sequences assigned to this particular cluster, SEQ ID NO:315 and SEQ ID NO:316, both displayed this differential expression, suggesting that the two sequences are likely associated with a single transcript.

SEQ ID NO:315 and SEQ ID NO:316 were then used as query sequences to search for homologous sequences in GenBank as described in Examples 1 and 2. SEQ ID NO: 315 displayed identity to the GenBank entry H72034 (SEQ ID NO:317) and SEQ ID NO:316 displayed identity to GenBank entry AA707002 (SEQ ID NO:318). SEQ ID NO:315 displays striking identity to the 3' end of SEQ ID NO:317 (See Figures 1A and 1B), while SEQ ID NO:316 displays striking identity to the 5' end of SEQ ID NO:318 (See Figure 2). Clones of H72034 and AA707002 were ordered

from the I.M.A.G.E. Consortium at the Lawrence Livermore National Laboratories (Livermore, California) for further studies.

Restriction Mapping of Clones H72034 and AA707002

5 The newly identified sequences were digested with a number of different restriction endonucleases to construct a restriction map of each of the clones. An appropriate amount of each clone, SEQ ID NO:317 or SEQ ID NO:318, was digested with various enzymes, and the restriction fragments identified as follows:

SEQ ID NO:317

		Enzyme	#Cuts	Positions							
5		AluI	5	331	1029	1422	1595	1977			
		BamHI	2	1836	2089						
		BstEII	1	936							
		BstXI	1	1033							
		HaeIII	12	145	300	453	497	582	780		
10	968			1102	1536	1561	1722	1981	2062		
		Hinfi	12	5	154	205	325	397	473	610	820
					1295	1426	2066				
		KpnI	1	1938							
		MspI	6	78	739	1098	2038	2077	2093		
15		NcoI	2	2013	2058						
		PstI	1	1501							
		PvuII	2	331	1422						
		Sau3AI	6	1270	1813	1819	1836	1894	2089		
		SphI	1	1870							
20		XhoI	1	1413							

SEQ ID NO:318

		Enzyme	#Cuts	Positions							
25	1214	AluI	9	19	245	367	553	586	874	904	996
		BamHI	1	407							
		BglI	1	1056							
		BglIII	1	475							
		BstEI	1	1108							
30	915	HaeIII	10	153	348	485	867	518	628	780	867
					1016	1312					
		HindIII	2	243	872						
		Hinfl	1	1353							
		KpnI	1	132							
35		MspI	2	1196	1261						
		PstI	1	823							
		PvuII	1	996							
		Sau3AI	7	66	407	475	504	750	850	1024	

The restriction maps based on the identified sites can be used to determine the position of each clone relative to the genomic sequences, and to confirm the 5'-3' orientation of the clones.

Amplification and Purification of Transcript

A transcript in this region upregulated in low metastatic cancers which contain sequences from SEQ ID NOS: 315-318 is identified using a technique such as polymerase chain reaction (PCR) amplification. Based on the sequences identified and the original sequences of the cluster, primers can be designed to isolate the full length cDNA from a library constructed from the breast cancer cell line with low metastatic potential.

A cDNA template for use in the amplification reaction is generated from total RNA isolated from the high metastatic breast cell line. RNA is reverse transcribed using oligo-dT primer to generate first strand cDNA. cDNA is synthesized by denaturing 3 µl of total RNA, 2 µl oligo-dT primer at 20 µM, and 5 µl DEPC water for 8 minutes at 65°C followed by reverse transcription at 52°C for 1 hour in a reaction containing the denatured RNA/primer plus 4 µl 1 5X cDNA buffer (GibcoBRL), 1 µl 0.1 M dithiothreitol, 1 µl 40 U/l RNaseOUT (GibcoBRL), 1 µl DEPC water, 2 µl 10 mM dNTP (GibcoBRL), and 1 µl 15 U/l Thermoscript reverse transcriptase (GibcoBRL). The reaction was terminated by a 5-min incubation at 85°C, and the RNA was removed by 1 µl 2 U/l RNase H at 37°C for thirty minutes.

Based on the determined orientation of the clones, primers are designed to amplify a full-length clone corresponding to the differentially expressed transcript in this region. Forward primers that are used to amplify the full-length clone are taken from the 5' end of SEQ ID NO:17 as follows:

F1 5'- TGGGATATAGTCTCGTGGTGCG -3' (SEQ ID NO:319)

F2 5'- TGATTCGATGTCATCAGTCCCG-3' (SEQ ID NO:320)

Primer F1 is taken from residues 51-62 of SEQ ID NO: 317, and primer F2 is taken from residues 212-233 Of SEQ ID NO:17. Both forward primers are near the 5' end of this sequence.

Reverse Primers are designed using sequences complementary to the 3' end of clone 10154-3 as follows:

R1 5'- TGTGTCACAGCCAGACATGAGC (SEQ ID NO:321)

R2 5' – TGCAAACATACACAGGGACCG (SEQ ID NO:322)

Primer R1 is based on residues 573-552 of SEQ ID NO:318, and R2 is based on residues 399-379 of SEQ ID NO:318.

PCR is performed using a 5µl aliquot of the first strand cDNA synthesis reaction, and a primer pair, e.g., F1 and R1, F1 and R2, F2 and R1, or F2 and R2. An open reading frame is amplified using 2 µl of the reverse transcription product as template in a PCR reaction containing 5 µl of 10x PCR buffer (GibcoBRL), 1 µl 50 mM Mg₂SO₄, 1 µl 10 mM dNTP, 1 µl F1 or F2 primer, 1 µl R1 primer, 2.5 U High Fidelity Platinum Taq DNA polymerase (GibcoBRL), and water to 50 µl. The molecule is amplified using 30 rounds of amplification in a thermal cycler at the following temperatures: 1 minute at 95°C; 1 minute at 55°C and 2 minutes at 72°C. The 30 cycles was followed by a 10 minute extension at 72°C.

Following amplification of the sequences, the PCR products are loaded on a 1% TEA gel and subjected to gel purification. One or more bands can be isolated from the gel and the DNA was purified using a QIAquick® Gel Extraction Kit (Qiagen, Valencia, CA). The purified fragment was cloned into a bacterial vector and transformed into the bacterial strain DH5α. Following cloning of the purified fragment(s), the DNA can be isolated and sequenced to confirm that a band corresponds to a transcript from this genetic region.

The reactions are carried out with two different 5' and 3' primers to increase the likelihood that the reaction will yield an amplification product. Other primers may also be designed from the predicted 5' and/or 3' end of the sequence, as will be apparent to one skilled in the art upon reading this disclosure, and thus other primers may be designed from the general region of SEQ ID NOS:317 and 318 that may yield better results than the disclosed primers.

In order to obtain additional sequences 5' to the end of a partial cDNA, 5' rapid amplification of cDNA ends (RACE) can be performed to ensure that the entire transcript has been identified. See *PCR Protocols: A Guide to Methods and Applications*, (1990) Academic Press, Inc. Following isolation of a cDNA using the F1-R1 or F2-R1 primer pairs, additional primers can be designed to perform RACE. The primers can be designed from the sequence of 10154-1 as follows:

5'-TTTAGCAGCACTAATGACTGTGGC-3' (SEQ ID NO:323)

5'-CGCCGTGAATTACTGTGGATGG-3' (SEQ ID NO:324)

The two RACE primers are designed based residues 286-263 and 396-375 of SEQ ID NO:317, respectively.

These sequences can be used to obtain any transcript sequences 5' to the amplification products obtained using the PCR protocol described above.

Northern Analysis

Other techniques can be used for confirming differential expression of the full-length transcript. For example, a Northern Blot can be used to verify differential expression of SEQ ID

NOS:317 and 318 in a breast cancer cells with low metastatic potential compared to breast cancer cells with high metastatic potential. Northern analysis can be accomplished by methods well-known in the art. Briefly, RNA is individually isolated from breast cancer cells having high metastatic potential and breast cancer cells having low metastatic potential, *e.g.*, a product such as RNeasy Mini Kits (Qiagen, CA) or NucleoSpin® RNA II Kit (Clontech, Palo Alto, CA). The isolated RNA samples are For Northern analysis, RNA isolated from the cells was electrophoresed on a denaturing formaldehyde agarose gel and transferred onto a membrane such as a supported nitrocellulose membrane (Schleicher & Schuell).

Rapid-Hyb buffer (Amersham Life Science, Little Chalfont, England) with 5 mg/ml denatured single stranded sperm DNA is pre-warmed to 65°C and the RNA blots are pre-hybridized in the buffer with shaking at 65°C for 30 minutes. Gene-specific DNA probes (50 ng per reaction) labeled with [α -³²P]dCTP (3000Ci/mmol, Amersham Pharmacia Biotech Inc., Piscataway, NJ) (Prime-It RmT Kit, Stratagene, La Jolla, CA) and purified with ProbeQuant™ G-50 Micro Columns (Amersham Pharmacia Biotech Inc.) are added and hybridized to the blots with shaking at 65°C for overnight. The blots are washed in 2x SSC, 0.1%(w/v) SDS at room temperature for 20 minutes, twice in 1x SSC, 0.1%(w/v) SDS at 65°C for 15 minutes, then exposed to Hyperfilms (Amersham Life Science).

Example 6: Identification of Differentially Expressed Genes by Array Analysis with Patient Tissue Samples

Differentially expressed genes corresponding to the polynucleotides described herein were also identified by microarray hybridization analysis using materials obtained from patient tissue samples. The biological materials used in these experiments are described below.

Source of patient tissue samples

Normal and cancerous tissues were collected from patients using laser capture microdissection (LCM) techniques, which techniques are well known in the art (see, *e.g.*, Ohyama *et al.* (2000) *Biotechniques* 29:530-6; Curran *et al.* (2000) *Mol. Pathol.* 53:64-8; Suarez-Quian *et al.* (1999) *Biotechniques* 26:328-35; Simone *et al.* (1998) *Trends Genet* 14:272-6; Conia *et al.* (1997) *J. Clin. Lab. Anal.* 11:28-38; Emmert-Buck *et al.* (1996) *Science* 274:998-1001). Table 8 (inserted following the last page of the Examples) provides information about each patient from which the samples were isolated, including: the Patient ID and Path ReportID, numbers assigned to the patient and the pathology reports for identification purposes; the anatomical location of the tumor (AnatomicalLoc); The Primary Tumor Size; the Primary Tumor Grade; the Histopathologic Grade; a description of local sites to which the tumor had invaded (Local Invasion); the presence of lymph node metastases (Lymph Node Metastasis); incidence of lymph node metastases (provided as number of lymph nodes positive for metastasis over the number of lymph nodes examined)

(Incidence Lymphnode Metastasis); the Regional Lymphnode Grade; the identification or detection of metastases to sites distant to the tumor and their location (Distant Met & Loc); a description of the distant metastases (Description Distant Met); the grade of distant metastasis (Distant Met Grade); and general comments about the patient or the tumor (Comments). Adenoma was not described in any of the patients. ; adenoma dysplasia (described as hyperplasia by the pathologist) was described in Patient ID No. 695. Extranodal extensions were described in two patients, Patient ID Nos. 784 and 791. Lymphovascular invasion was described in seven patients, Patient ID Nos. 128, 278, 517, 534, 784, 786, and 791.. Crohn's-like infiltrates were described in seven patients, Patient ID Nos. 52, 264, 268, 392, 393, 784, and 791.

Source of polynucleotides on arrays

Polynucleotides on arrays

Polynucleotides spotted on the arrays were generated by PCR amplification of clones derived from cDNA libraries. The clones used for amplification were either the clones from which the sequences described herein (SEQ ID NOS:1-316) were derived, or are clones having inserts with significant polynucleotide sequence overlap with the sequences described herein (SEQ ID NO:1-316) as determined by BLAST2 homology searching.

Microarray Design

Each array used in the examples below had an identical spatial layout and control spot set. Each microarray was divided into two areas, each area having an array with, on each half, twelve groupings of 32 x 12 spots for a total of about 9,216 spots on each array. The two areas are spotted identically which provide for at least two duplicates of each clone per array. Spotting was accomplished using PCR amplified products from 0.5kb to 2.0 kb and spotted using a Molecular Dynamics Gen III spotter according to the manufacturer's recommendations. The first row of each of the 24 regions on the array had about 32 control spots, including 4 negative control spots and 8 test polynucleotides.

The test polynucleotides were spiked into each sample before the labeling reaction with a range of concentrations from 2-600 pg/slide and ratios of 1:1 . For each array design, two slides were hybridized with the test samples reverse-labeled in the labeling reaction. This provided for about 4 duplicate measurements for each clone, two of one color and two of the other, for each sample.

Microarray analysis

cDNA probes were prepared from total RNA isolated from the patient cells described in above (Table 8). Since LCM provides for the isolation of specific cell types to provide a substantially homogenous cell sample, this provided for a similarly pure RNA sample.

Total RNA was first reverse transcribed into cDNA using a primer containing a T7 RNA polymerase promoter, followed by second strand DNA synthesis. cDNA was then transcribed *in*

vitro to produce antisense RNA using the T7 promoter-mediated expression (see, *e.g.*, Luo *et al.* (1999) *Nature Med* 5:117-122), and the antisense RNA was then converted into cDNA. The second set of cDNAs were again transcribed *in vitro*, using the T7 promoter, to provide antisense RNA. Optionally, the RNA was again converted into cDNA, allowing for up to a third round of T7-mediated amplification to produce more antisense RNA. Thus the procedure provided for two or three rounds of *in vitro* transcription to produce the final RNA used for fluorescent labeling. Fluorescent probes were generated by first adding control RNA to the antisense RNA mix, and producing fluorescently labeled cDNA from the RNA starting material. Fluorescently labeled cDNAs prepared from the tumor RNA sample were compared to fluorescently labeled cDNAs prepared from normal cell RNA sample. For example, the cDNA probes from the normal cells were labeled with Cy3 fluorescent dye (green) and the cDNA probes prepared from the tumor cells were labeled with Cy5 fluorescent dye (red).

The differential expression assay was performed by mixing equal amounts of probes from tumor cells and normal cells of the same patient. The arrays were prehybridized by incubation for about 2 hrs at 60°C in 5X SSC/0.2% SDS/1 mM EDTA, and then washed three times in water and twice in isopropanol. Following prehybridization of the array, the probe mixture was then hybridized to the array under conditions of high stringency (overnight at 42°C in 50% formamide, 5X SSC, and 0.2% SDS. After hybridization, the array was washed at 55°C three times as follows: 1) first wash in 1X SSC/0.2% SDS; 2) second wash in 0.1X SSC/0.2% SDS; and 3) third wash in 0.1X SSC.

The arrays were then scanned for green and red fluorescence using a Molecular Dynamics Generation III dual color laser-scanner/detector. The images were processed using BioDiscovery Autogene software, and the data from each scan set normalized to provide for a ratio of expression relative to normal. Data from the microarray experiments was analyzed according to the algorithms described in U.S. application serial no. 60/252,358, filed November 20, 2000, by E.J. Moler, M.A. Boyle, and F.M. Randazzo, and entitled "Precision and accuracy in cDNA microarray data," which application is specifically incorporated herein by reference.

The experiment was repeated, this time labeling the two probes with the opposite color in order to perform the assay in both "color directions." Each experiment was sometimes repeated with two more slides (one in each color direction). The level fluorescence for each sequence on the array expressed as a ratio of the geometric mean of 8 replicate spots/genes from the four arrays or 4 replicate spots/gene from 2 arrays or some other permutation. The data were normalized using the spiked positive controls present in each duplicated area, and the precision of this normalization was included in the final determination of the significance of each differential. The fluorescent intensity of each spot was also compared to the negative controls in each duplicated area to determine which spots have detected significant expression levels in each sample.

A statistical analysis of the fluorescent intensities was applied to each set of duplicate spots to assess the precision and significance of each differential measurement, resulting in a p-value testing the null hypothesis that there is no differential in the expression level between the tumor and normal samples of each patient. For initial analysis of the microarrays, the hypothesis was accepted if $p > 10^{-3}$, and the differential ratio was set to 1.000 for those spots. All other spots have a significant difference in expression between the tumor and normal sample. If the tumor sample has detectable expression and the normal does not, the ratio is truncated at 1000 since the value for expression in the normal sample would be zero, and the ratio would not be a mathematically useful value (e.g., infinity). If the normal sample has detectable expression and the tumor does not, the ratio is truncated to 0.001, since the value for expression in the tumor sample would be zero and the ratio would not be a mathematically useful value. These latter two situations are referred to herein as "on/off." Database tables were populated using a 95% confidence level ($p > 0.05$).

Table 9 below summarize the results of the differential expression analysis. Each table provides: the SEQ ID NO of the polynucleotide corresponding to the polynucleotide on the spot on the array; the Spot ID (an identifier assigned to the spot so as to distinguish it from spots on the same and different arrays), the number of patients for whom there was information obtained from the array (Num Ratios), and the percentage of patients in which expression was detected at greater than or equal to a two-fold increase ($\geq 2x$), greater than or equal to a five-fold increase ($\geq 5x$), or less than or equal to a 1/2 -fold decrease ($\leq \text{half}$) relative to matched normal control tissue.

In general, a polynucleotide is said to represent a significantly differentially expressed gene between two samples when there is detectable levels of expression in at least one sample and the ratio value is greater than at least about 1.2 fold, preferably greater than at least about 1.5 fold, more preferably greater than at least about 2 fold, where the ratio value is calculated using the method described above.

A differential expression ratio of 1 indicates that the expression level of the gene in the tumor cell was not statistically different from expression of that gene in normal colon cells of the same patient. A differential expression ratio significantly greater than 1 in cancerous colon cells relative to normal colon cells indicates that the gene is increased in expression in cancerous cells relative to normal cells, indicating that the gene plays a role in the development of the cancerous phenotype, and may be involved in promoting metastasis of the cell. Detection of gene products from such genes can provide an indicator that the cell is cancerous, and may provide a therapeutic and/or diagnostic target.

Likewise, a differential expression ratio significantly less than 1 in cancerous colon cells relative to normal colon cells indicates that, for example, the gene is involved in suppression of the cancerous phenotype. Increasing activity of the gene product encoded by such a gene, or replacing such activity, can provide the basis for chemotherapy. Such gene can also serve as markers of

cancerous cells, e.g., the absence or decreased presence of the gene product in a colon cell relative to a normal colon cell indicates that the cell may be cancerous.

Table 9.

SEQ ID NO:	SpotID	Num Ratios	$\geq 2x$	$\geq 5x$	$\leq \text{halfx}$
8	579	33	87.88	39.39	3.03
12	22300	33	33.33	18.18	6.06
26	21886	33	33.33	0.00	3.03
64	9487	33	33.33	12.12	3.03
248	28179	28	32.14	0.00	0.00
253	28179	28	32.14	0.00	0.00
272	28179	28	32.14	0.00	0.00
292	9111	33	33.33	18.18	3.03
295	19980	33	33.33	6.06	0.00
309	23993	33	42.42	3.03	3.03

Deposit Information. The following materials were deposited with the American Type Culture Collection (CMCC = Chiron Master Culture Collection).

Table 5. Cell Lines Deposited with ATCC

Cell Line	Deposit Date	ATCC Accession No.	CMCC Accession No.
KM12L4-A	March 19, 1998	CRL-12496	11606
Km12C	May 15, 1998	CRL-12533	11611
MDA-MB-231	May 15, 1998	CRL-12532	10583
MCF-7	October 9, 1998	CRL-12584	10377

In addition, pools of selected clones, as well as libraries containing specific clones, were assigned an "ES" number (internal reference) and deposited with the ATCC. Table 6 below provides the ATCC Accession Nos. of the ES deposits, all of which were deposited on or before May 13, 1999. The names of the clones contained within each of these deposits are provided in the Table 7 (inserted before the claims).

Table 6: Pools of Clones and Libraries Deposited with ATCC on or before March 28, 2000

Cell Line	CMCC	ATCC
ES75	5140	PTA-1102
ES76	5141	PTA-1103
ES77	5142	PTA-1104
ES78	5143	PTA-1105
ES79	5144	PTA-1106
ES80	5145	PTA-1107
ES81	5146	PTA-1108
ES82	5147	PTA-1109
ES83	5148	PTA-1110
ES84	5149	PTA-1111

The deposits described herein are provided merely as convenience to those of skill in the art, and is not an admission that a deposit is required under 35 U.S.C. §112. The sequence of the polynucleotides contained within the deposited material, as well as the amino acid sequence of the polypeptides encoded thereby, are incorporated herein by reference and are controlling in the event
5 of any conflict with the written description of sequences herein. A license may be required to make, use, or sell the deposited material, and no such license is granted hereby.

Retrieval of Individual Clones from Deposit of Pooled Clones. Where the ATCC deposit is composed of a pool of cDNA clones or a library of cDNA clones, the deposit was prepared by first
10 transfecting each of the clones into separate bacterial cells. The clones in the pool or library were then deposited as a pool of equal mixtures in the composite deposit. Particular clones can be obtained from the composite deposit using methods well known in the art. For example, a bacterial cell containing a particular clone can be identified by isolating single colonies, and identifying colonies containing the specific clone through standard colony hybridization techniques, using an
15 oligonucleotide probe or probes designed to specifically hybridize to a sequence of the clone insert (*e.g.*, a probe based upon unmasked sequence of the encoded polynucleotide having the indicated SEQ ID NO). The probe should be designed to have a T_m of approximately 80°C (assuming 2°C for each A or T and 4°C for each G or C). Positive colonies can then be picked, grown in culture, and the recombinant clone isolated. Alternatively, probes designed in this manner can be used to
20 PCR to isolate a nucleic acid molecule from the pooled clones according to methods well known in the art, *e.g.*, by purifying the cDNA from the deposited culture pool, and using the probes in PCR reactions to produce an amplified product having the corresponding desired polynucleotide sequence.

Those skilled in the art will recognize, or be able to ascertain, using not more than routine
25 experimentation, many equivalents to the specific embodiments of the invention described herein. Such specific embodiments and equivalents are intended to be encompassed by the following claims.

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and
30 individually indicated to be incorporated by reference. The citation of any publication is for its disclosure prior to the filing date and should not be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it is readily apparent to those of ordinary skill
35 in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
1	819545	RTA22200265F.k.06.1.P.Seq	F	M00064554D:A03	CH22PRC
2	377944	RTA22200251F.j.02.1.P.Seq	F	M00063482A:A08	CH21PRN
3	818497	RTA22200252F.a.13.1.P.Seq	F	M00063514C:D03	CH21PRN
4	819498	RTA22200252F.n.05.1.P.Seq	F	M00063638C:G12	CH21PRN
5	455465	RTA22200264F.e.16.1.P.Seq	F	M00064454A:H10	CH22PRC
6	819069	RTA22200255F.f.01.1.P.Seq	F	M00063940D:F09	CH21PRN
7	672003	RTA22200265F.b.09.1.P.Seq	F	M00064517C:F11	CH22PRC
8	728115	RTA22200253F.o.24.1.P.Seq	F	M00063838B:G08	CH21PRN
9	372700	RTA22200260F.b.20.1.P.Seq	F	M00063580C:A06	CH22PRC
10	818056	RTA22200266F.c.13.1.P.Seq	F	M00064593D:C01	CH22PRC
11	818497	RTA22200255F.a.17.1.P.Seq	F	M00063920D:H02	CH21PRN
12	729832	RTA22200267F.l.21.1.P.Seq	F	M00064714A:G03	CH22PRC
13	505514	RTA22200251F.b.21.1.P.Seq	F	M00063158A:A01	CH21PRN
14	376488	RTA22200254F.c.05.1.P.Seq	F	M00063852B:D08	CH21PRN
15	376488	RTA22200260F.b.09.1.P.Seq	F	M00063578C:A06	CH22PRC
16	748572	RTA22200254F.c.07.1.P.Seq	F	M00063852D:F07	CH21PRN
17	549934	RTA22200253F.k.18.1.P.Seq	F	M00063801B:D04	CH21PRN
18	819069	RTA22200255F.e.24.1.P.Seq	F	M00063940D:F09	CH21PRN
19	817618	RTA22200253F.n.16.1.P.Seq	F	M00063828D:E05	CH21PRN
20	124396	RTA22200263F.a.11.2.P.Seq	F	M00064375B:G07	CH22PRC
21	404375	RTA22200260F.m.08.1.P.Seq	F	M00063967D:G02	CH22PRC
22	391820	RTA22200261F.f.02.1.P.Seq	F	M00064000B:C03	CH22PRC
23	672003	RTA22200267F.i.06.1.P.Seq	F	M00064693D:F08	CH22PRC
24	830620	RTA22200263F.n.09.1.P.Seq	F	M00064424B:C12	CH22PRC
25	450399	RTA22200251F.f.23.1.P.Seq	F	M00063467D:H07	CH21PRN
26	450982	RTA22200261F.n.18.1.P.Seq	F	M00064307B:G02	CH22PRC
27	819894	RTA22200264F.h.18.1.P.Seq	F	M00064467B:D06	CH22PRC
28	379302	RTA22200257F.j.02.3.P.Seq	F	M00064178C:C04	CH21PRN
29	379746	RTA22200256F.e.16.1.P.Seq	F	M00064086C:E01	CH21PRN
30	124863	RTA22200265F.m.06.1.P.Seq	F	M00064564A:C02	CH22PRC
31	379154	RTA22200257F.c.11.1.P.Seq	F	M00064151B:C07	CH21PRN
32	830620	RTA22200262F.l.23.1.P.Seq	F	M00064358C:D09	CH22PRC
33	389409	RTA22200266F.l.24.1.P.Seq	F	M00064631A:C07	CH22PRC
34	397284	RTA22200262F.i.22.1.P.Seq	F	M00064346C:B09	CH22PRC
35	819440	RTA22200264F.e.19.1.P.Seq	F	M00064454C:B06	CH22PRC
36	389409	RTA22200266F.m.01.1.P.Seq	F	M00064631A:C07	CH22PRC
37	518848	RTA22200265F.n.15.1.P.Seq	F	M00064571C:C04	CH22PRC
38	830620	RTA22200263F.a.21.1.P.Seq	F	M00064376A:A05	CH22PRC
39	379154	RTA22200256F.f.20.1.P.Seq	F	M00064090D:D09	CH21PRN
40	818544	RTA22200256F.h.04.1.P.Seq	F	M00064105B:A03	CH21PRN
41	817375	RTA22200251F.a.15.1.P.Seq	F	M00063152C:B07	CH21PRN

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
42	455264	RTA22200259F.e.23.1.P.Seq	F	M00063539C:C11	CH22PRC
43	817503	RTA22200266F.k.11.1.P.Seq	F	M00064624D:C09	CH22PRC
44	377696	RTA22200256F.d.21.1.P.Seq	F	M00064082D:D10	CH21PRN
45	375596	RTA22200261F.h.10.1.P.Seq	F	M00064009A:C01	CH22PRC
46	817689	RTA22200263F.h.05.1.P.Seq	F	M00064399A:E01	CH22PRC
47	831867	RTA22200262F.i.15.2.P.Seq	F	M00064345A:A03	CH22PRC
48	830085	RTA22200261F.k.14.1.P.Seq	F	M00064293D:B12	CH22PRC
49	389627	RTA22200264F.c.10.1.P.Seq	F	M00064447B:C06	CH22PRC
50	397284	RTA22200259F.k.09.1.P.Seq	F	M00063555B:D01	CH22PRC
51	380063	RTA22200261F.j.02.1.P.Seq	F	M00064014D:H05	CH22PRC
52	830931	RTA22200266F.m.23.1.P.Seq	F	M00064633C:A03	CH22PRC
53	819321	RTA22200257F.l.03.3.P.Seq	F	M00064194C:D02	CH21PRN
54	475587	RTA22200261F.c.01.1.P.Seq	F	M00063990A:D05	CH22PRC
55	819046	RTA22200255F.a.18.1.P.Seq	F	M00063920D:H05	CH21PRN
56	817477	RTA22200253F.g.21.1.P.Seq	F	M00063784A:H12	CH21PRN
57	475587	RTA22200261F.b.24.1.P.Seq	F	M00063990A:D05	CH22PRC
58	728115	RTA22200253F.p.01.1.P.Seq	F	M00063838B:G08	CH21PRN
59	389627	RTA22200260F.i.24.1.P.Seq	F	M00063957A:E02	CH22PRC
60	403453	RTA22200256F.i.24.1.P.Seq	F	M00064113B:C04	CH21PRN
61	508525	RTA22200255F.d.10.1.P.Seq	F	M00063931B:F07	CH21PRN
62	819525	RTA22200261F.n.20.1.P.Seq	F	M00064307C:G03	CH22PRC
63	817618	RTA22200255F.i.03.1.P.Seq	F	M00064025D:H12	CH21PRN
64	819403	RTA22200254F.h.14.1.P.Seq	F	M00063888D:D05	CH21PRN
65	553242	RTA22200254F.g.20.1.P.Seq	F	M00063886A:B06	CH21PRN
66	817417	RTA22200255F.a.10.1.P.Seq	F	M00063919C:E07	CH21PRN
67	817618	RTA22200252F.f.13.1.P.Seq	F	M00063604A:B11	CH21PRN
68	611440	RTA22200262F.e.04.2.P.Seq	F	M00064328B:H09	CH22PRC
69	817375	RTA22200260F.m.06.1.P.Seq	F	M00063967C:A12	CH22PRC
70	213577	RTA22200255F.i.23.1.P.Seq	F	M00064033C:C11	CH21PRN
71	820061	RTA22200265F.p.10.1.P.Seq	F	M00064579D:E11	CH22PRC
72	455264	RTA22200259F.m.06.1.P.Seq	F	M00063559D:G03	CH22PRC
73	455264	RTA22200255F.o.23.1.P.Seq	F	M00064059A:C11	CH21PRN
74	380331	RTA22200255F.b.19.1.P.Seq	F	M00063926A:H04	CH21PRN
75	380331	RTA22200252F.b.19.1.P.Seq	F	M00063518D:A01	CH21PRN
76	817455	RTA22200267F.o.01.1.P.Seq	F	M00064723D:H03	CH22PRC
77	423967	RTA22200252F.a.20.1.P.Seq	F	M00063515B:H02	CH21PRN
78	220584	RTA22200261F.m.14.1.P.Seq	F	M00064302A:D10	CH22PRC
79	817688	RTA22200251F.e.20.1.P.Seq	F	M00063462D:D07	CH21PRN
80	549934	RTA22200253F.n.10.1.P.Seq	F	M00063826A:D03	CH21PRN
81	819149	RTA22200255F.e.16.1.P.Seq	F	M00063938B:H07	CH21PRN
82	817455	RTA22200267F.n.24.1.P.Seq	F	M00064723D:H03	CH22PRC

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
83	377696	RTA22200251F.j.03.1.P.Seq	F	M00063482A:F07	CH21PRN
84	830146	RTA22200260F.b.07.1.P.Seq	F	M00063578B:E02	CH22PRC
85	194490	RTA22200264F.l.07.1.P.Seq	F	M00064481C:F03	CH22PRC
86	819460	RTA22200257F.m.15.3.P.Seq	F	M00064200D:E08	CH21PRN
87	819018	RTA22200257F.p.01.3.P.Seq	F	M00064212D:E04	CH21PRN
88	830620	RTA22200259F.p.24.1.P.Seq	F	M00063571B:G03	CH22PRC
89	141079	RTA22200262F.k.19.1.P.Seq	F	M00064354A:A10	CH22PRC
90	376588	RTA22200256F.e.04.1.P.Seq	F	M00064083D:E05	CH21PRN
91	380604	RTA22200264F.g.05.1.P.Seq	F	M00064460C:B01	CH22PRC
92	413138	RTA22200260F.b.05.1.P.Seq	F	M00063577C:C02	CH22PRC
93	818544	RTA22200265F.e.12.1.P.Seq	F	M00064527A:H07	CH22PRC
94	647435	RTA22200257F.h.08.1.P.Seq	F	M00064172C:A02	CH21PRN
95	551785	RTA22200266F.c.09.1.P.Seq	F	M00064593A:A05	CH22PRC
96	17092	RTA22200261F.f.17.1.P.Seq	F	M00064002C:F06	CH22PRC
97	818326	RTA22200251F.i.06.1.P.Seq	F	M00063478C:D01	CH21PRN
98	377944	RTA22200262F.e.03.2.P.Seq	F	M00064328B:H04	CH22PRC
99	745559	RTA22200262F.m.04.1.P.Seq	F	M00064359B:H12	CH22PRC
100	818326	RTA22200265F.d.08.1.P.Seq	F	M00064524A:A09	CH22PRC
101	379879	RTA22200264F.b.23.1.P.Seq	F	M00064446A:D11	CH22PRC
102	819640	RTA22200257F.f.24.1.P.Seq	F	M00064165A:B12	CH21PRN
103	818326	RTA22200265F.a.14.1.P.Seq	F	M00064514D:F11	CH22PRC
104	243524	RTA22200265F.g.04.1.P.Seq	F	M00064532D:G06	CH22PRC
105	43995	RTA22200261F.l.02.1.P.Seq	F	M00064294D:F01	CH22PRC
106	597854	RTA22200262F.g.06.2.P.Seq	F	M00064337D:F01	CH22PRC
107	268290	RTA22200260F.p.14.1.P.Seq	F	M00063981D:A06	CH22PRC
108	818043	RTA22200256F.p.10.2.P.Seq	F	M00064138A:F11	CH21PRN
109	830930	RTA22200267F.b.03.1.P.Seq	F	M00064652B:D09	CH22PRC
110	389627	RTA22200260F.j.01.1.P.Seq	F	M00063957A:E02	CH22PRC
111	378730	RTA22200260F.i.07.1.P.Seq	F	M00063955C:F07	CH22PRC
112	819037	RTA22200260F.n.09.1.P.Seq	F	M00063972C:E10	CH22PRC
113	830397	RTA22200261F.g.14.1.P.Seq	F	M00064005D:A08	CH22PRC
114	450247	RTA22200261F.e.10.1.P.Seq	F	M00063998C:E09	CH22PRC
115	819273	RTA22200252F.b.09.1.P.Seq	F	M00063517A:A04	CH21PRN
116	587779	RTA22200257F.i.11.3.P.Seq	F	M00064175B:B09	CH21PRN
117	818639	RTA22200256F.j.09.1.P.Seq	F	M00064115B:E12	CH21PRN
118	615617	RTA22200261F.o.13.1.P.Seq	F	M00064309C:H09	CH22PRC
119	79309	RTA22200257F.j.13.3.P.Seq	F	M00064180A:G03	CH21PRN
120	748994	RTA22200261F.o.20.1.P.Seq	F	M00064310C:A10	CH22PRC
121	818682	RTA22200258F.h.07.1.P.Seq	F	M00064271B:D03	CH21PRN
122	373061	RTA22200253F.j.09.1.P.Seq	F	M00063795C:D09	CH21PRN
123	484413	RTA22200253F.g.09.1.P.Seq	F	M00063781B:B10	CH21PRN

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
124	819273	RTA22200258F.h.04.1.P.Seq	F	M00064270B:B03	CH21PRN
125	569532	RTA22200252F.h.18.1.P.Seq	F	M00063613D:C11	CH21PRN
126	170313	RTA22200255F.g.20.1.P.Seq	F	M00063949D:A05	CH21PRN
127	818682	RTA22200253F.p.14.1.P.Seq	F	M00063841A:B09	CH21PRN
128	377188	RTA22200255F.l.06.1.P.Seq	F	M00064043D:C09	CH21PRN
129	518848	RTA22200257F.j.22.3.P.Seq	F	M00064186C:B03	CH21PRN
130	45592	RTA22200259F.l.08.1.P.Seq	F	M00063557D:C07	CH22PRC
131	819273	RTA22200255F.n.19.1.P.Seq	F	M00064053C:G04	CH21PRN
132	397284	RTA22200251F.a.06.1.P.Seq	F	M00063151D:B10	CH21PRN
133	818326	RTA22200258F.e.14.1.P.Seq	F	M00064260C:E05	CH21PRN
134	819037	RTA22200251F.c.15.1.P.Seq	F	M00063452A:F08	CH21PRN
135	817417	RTA22200253F.m.14.1.P.Seq	F	M00063818C:A09	CH21PRN
136	819640	RTA22200254F.i.11.1.P.Seq	F	M00063891A:F11	CH21PRN
137	818771	RTA22200254F.i.19.1.P.Seq	F	M00063892B:G02	CH21PRN
138	389627	RTA22200254F.k.10.1.P.Seq	F	M00063898A:A10	CH21PRN
139	379067	RTA22200260F.e.20.1.P.Seq	F	M00063593A:D03	CH22PRC
140	818544	RTA22200251F.f.02.1.P.Seq	F	M00063463D:B05	CH21PRN
141	819440	RTA22200251F.j.22.1.P.Seq	F	M00063485A:E05	CH21PRN
142	817417	RTA22200251F.k.10.1.P.Seq	F	M00063487C:C02	CH21PRN
143	385307	RTA22200262F.k.11.1.P.Seq	F	M00064352C:H01	CH22PRC
144	611440	RTA22200263F.d.24.2.P.Seq	F	M00064386B:C02	CH22PRC
145	376056	RTA22200259F.e.16.1.P.Seq	F	M00063538D:B01	CH22PRC
146	611440	RTA22200263F.d.24.1.P.Seq	F	M00064386B:C02	CH22PRC
147	820061	RTA22200264F.f.09.1.P.Seq	F	M00064457D:C09	CH22PRC
148	617825	RTA22200264F.p.06.1.P.Seq	F	M00064508A:B09	CH22PRC
149	819440	RTA22200257F.h.17.1.P.Seq	F	M00064173B:E01	CH21PRN
150	819145	RTA22200266F.m.08.1.P.Seq	F	M00064631C:H11	CH22PRC
151	817653	RTA22200265F.p.07.1.P.Seq	F	M00064579A:C06	CH22PRC
152	611440	RTA22200263F.e.01.1.P.Seq	F	M00064386B:C02	CH22PRC
153	375958	RTA22200264F.j.22.1.P.Seq	F	M00064476D:C04	CH22PRC
154	611440	RTA22200257F.a.20.1.P.Seq	F	M00064144D:A07	CH21PRN
155	831049	RTA22200266F.o.13.1.P.Seq	F	M00064637B:F03	CH22PRC
156	818162	RTA22200266F.g.18.1.P.Seq	F	M00064610D:H01	CH22PRC
157	553200	RTA22200263F.p.02.1.P.Seq	F	M00064429D:B07	CH22PRC
158	139677	RTA22200254F.o.07.1.P.Seq	F	M00063910D:A12	CH21PRN
159	139677	RTA22200252F.c.11.1.P.Seq	F	M00063520D:E11	CH21PRN
160	397284	RTA22200262F.i.22.2.P.Seq	F	M00064346C:B09	CH22PRC
161	385810	RTA22200256F.m.04.2.P.Seq	F	M00064126C:F12	CH21PRN
162	404624	RTA22200261F.e.07.1.P.Seq	F	M00063997C:B12	CH22PRC
163	375958	RTA22200262F.b.14.2.P.Seq	F	M00064322C:A10	CH22PRC
164	616555	RTA22200265F.b.24.1.P.Seq	F	M00064520A:E04	CH22PRC

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
165	616555	RTA22200265F.c.01.1.P.Seq	F	M00064520A:E04	CH22PRC
166	295694	RTA22200260F.o.20.1.P.Seq	F	M00063978B:B06	CH22PRC
167	36113	RTA22200265F.e.06.1.P.Seq	F	M00064526D:F05	CH22PRC
168	831812	RTA22200263F.f.05.1.P.Seq	F	M00064390A:C05	CH22PRC
169	817653	RTA22200252F.g.23.1.P.Seq	F	M00063610D:C11	CH21PRN
170	397284	RTA22200252F.m.15.1.P.Seq	F	M00063636A:E01	CH21PRN
171	817979	RTA22200253F.p.15.1.P.Seq	F	M00063841A:E08	CH21PRN
172	817653	RTA22200255F.m.18.1.P.Seq	F	M00064048C:G12	CH21PRN
173	611440	RTA22200253F.f.03.1.P.Seq	F	M00063774A:D09	CH21PRN
174	386014	RTA22200261F.f.06.1.P.Seq	F	M00064001A:B03	CH22PRC
175	549981	RTA22200255F.b.10.1.P.Seq	F	M00063925B:F04	CH21PRN
176	193373	RTA22200255F.l.21.1.P.Seq	F	M00064046A:G02	CH21PRN
177	400619	RTA22200255F.g.14.1.P.Seq	F	M00063947D:D01	CH21PRN
178	831149	RTA22200261F.o.21.1.P.Seq	F	M00064310D:F03	CH22PRC
179	36113	RTA22200255F.d.16.1.P.Seq	F	M00063932D:G08	CH21PRN
180	817503	RTA22200253F.l.16.1.P.Seq	F	M00063805D:E05	CH21PRN
181	376588	RTA22200260F.i.11.1.P.Seq	F	M00063955D:F05	CH22PRC
182	141079	RTA22200252F.f.23.1.P.Seq	F	M00063606C:B04	CH21PRN
183	818063	RTA22200253F.p.04.1.P.Seq	F	M00063839A:F01	CH21PRN
184	455264	RTA22200253F.n.14.1.P.Seq	F	M00063828A:H12	CH21PRN
185	189234	RTA22200251F.f.17.1.P.Seq	F	M00063466C:C11	CH21PRN
186	295694	RTA22200265F.j.05.1.P.Seq	F	M00064550A:A07	CH22PRC
187	648679	RTA22200260F.f.06.1.P.Seq	F	M00063594B:H07	CH22PRC
188	830930	RTA22200264F.e.10.1.P.Seq	F	M00064452D:E11	CH22PRC
189	818497	RTA22200256F.d.07.1.P.Seq	F	M00064079C:A10	CH21PRN
190	373928	RTA22200256F.d.19.1.P.Seq	F	M00064082A:A08	CH21PRN
191	385307	RTA22200263F.j.12.1.P.Seq	F	M00064406B:H06	CH22PRC
192	403453	RTA22200266F.e.10.1.P.Seq	F	M00064601D:B05	CH22PRC
193	730318	RTA22200264F.c.09.1.P.Seq	F	M00064447B:A07	CH22PRC
194	44183	RTA22200271F.a.01.1.P.Seq	F	M00021929A:D03	CH03MAH
195	373928	RTA22200255F.d.22.1.P.Seq	F	M00063934B:E04	CH21PRN
196	404624	RTA22200255F.d.23.1.P.Seq	F	M00063934C:C10	CH21PRN
197	403173	RTA22200253F.a.21.1.P.Seq	F	M00063685A:C02	CH21PRN
198	372700	RTA22200253F.c.06.1.P.Seq	F	M00063689D:E12	CH21PRN
199	374343	RTA22200261F.h.04.1.P.Seq	F	M00064008A:B01	CH22PRC
200	597854	RTA22200255F.j.03.1.P.Seq	F	M00064033D:B01	CH21PRN
201	817417	RTA22200255F.a.23.1.P.Seq	F	M00063922B:A12	CH21PRN
202	818497	RTA22200257F.k.05.3.P.Seq	F	M00064188B:G08	CH21PRN
203	377696	RTA22200255F.f.15.1.P.Seq	F	M00063943B:G12	CH21PRN
204	379105	RTA22200252F.n.19.1.P.Seq	F	M00063642B:A08	CH21PRN
205	831188	RTA22200267F.o.02.1.P.Seq	F	M00064723D:H11	CH22PRC

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
206	376056	RTA22200253F.m.09.1.P.Seq	F	M00063810C:E03	CH21PRN
207	124863	RTA22200255F.n.15.1.P.Seq	F	M00064053B:D09	CH21PRN
208	376056	RTA22200254F.i.03.1.P.Seq	F	M00063890A:F11	CH21PRN
209	831812	RTA22200266F.j.10.1.P.Seq	F	M00064620C:D01	CH22PRC
210	141079	RTA22200260F.i.14.1.P.Seq	F	M00063956A:F05	CH22PRC
211	19148	RTA22200265F.o.18.1.P.Seq	F	M00064577C:B12	CH22PRC
212	124396	RTA22200252F.a.14.1.P.Seq	F	M00063514C:E08	CH21PRN
213	831026	RTA22200265F.c.03.1.P.Seq	F	M00064520A:F08	CH22PRC
214	819037	RTA22200263F.i.23.1.P.Seq	F	M00064405B:C04	CH22PRC
215	380207	RTA22200263F.i.19.1.P.Seq	F	M00064404C:G05	CH22PRC
216	819460	RTA22200255F.c.13.1.P.Seq	F	M00063928A:G09	CH21PRN
217	379067	RTA22200253F.g.23.1.P.Seq	F	M00063784C:E10	CH21PRN
218	403173	RTA22200252F.p.23.1.P.Seq	F	M00063682A:C04	CH21PRN
219	3856	RTA22200269F.a.05.1.P.Seq	F	M00003773D:H02	CH01COH
220	378551	RTA22200263F.d.17.1.P.Seq	F	M00064385D:C11	CH22PRC
221	456089	RTA22200272F.a.09.1.P.Seq	F	M00043134A:A05	CH19COP
222	549981	RTA22200267F.a.22.1.P.Seq	F	M00064650B:B07	CH22PRC
223	378551	RTA22200265F.m.21.1.P.Seq	F	M00064568A:H06	CH22PRC
224	819201	RTA22200256F.n.23.2.P.Seq	F	M00064132B:B07	CH21PRN
225	374826	RTA22200251F.c.20.1.P.Seq	F	M00063453B:F08	CH21PRN
226	389409	RTA22200253F.l.23.1.P.Seq	F	M00063807A:D12	CH21PRN
227	819149	RTA22200260F.a.17.1.P.Seq	F	M00063575B:G02	CH22PRC
228	389409	RTA22200255F.e.18.1.P.Seq	F	M00063939C:D06	CH21PRN
229	818165	RTA22200254F.h.15.1.P.Seq	F	M00063888D:F02	CH21PRN
230	817757	RTA22200252F.i.15.1.P.Seq	F	M00063617D:F09	CH21PRN
231	553242	RTA22200263F.i.20.1.P.Seq	F	M00064404D:A06	CH22PRC
232	385615	RTA22200265F.b.08.1.P.Seq	F	M00064517B:F10	CH22PRC
233	819102	RTA22200258F.h.19.1.P.Seq	F	M00064272C:G01	CH21PRN
234	817757	RTA22200255F.o.16.1.P.Seq	F	M00064057C:H10	CH21PRN
235	385615	RTA22200265F.b.07.1.P.Seq	F	M00064517B:F04	CH22PRC
236	385615	RTA22200253F.l.06.1.P.Seq	F	M00063804C:A11	CH21PRN
237	827355	RTA22200266F.n.23.1.P.Seq	F	M00064636B:A04	CH22PRC
238	817629	RTA22200259F.a.13.1.P.Seq	F	M00063165A:C09	CH22PRC
239	817514	RTA22200260F.h.02.1.P.Seq	F	M00063600C:C09	CH22PRC
240	817514	RTA22200252F.p.21.1.P.Seq	F	M00063681B:C02	CH21PRN
241	680563	RTA22200265F.f.13.1.P.Seq	F	M00064530B:H02	CH22PRC
242	827355	RTA22200255F.e.20.1.P.Seq	F	M00063939C:H01	CH21PRN
243	377286	RTA22200254F.a.04.1.P.Seq	F	M00063843B:D07	CH21PRN
244	680563	RTA22200258F.g.18.1.P.Seq	F	M00064268D:G03	CH21PRN
245	819156	RTA22200255F.h.06.1.P.Seq	F	M00064021D:H01	CH21PRN
246	220584	RTA22200261F.f.22.1.P.Seq	F	M00064003B:C10	CH22PRC

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
247	616555	RTA22200263F.o.12.1.P.Seq	F	M00064428B:A12	CH22PRC
248	819498	RTA22200254F.o.14.1.P.Seq	F	M00063912A:D06	CH21PRN
249	817508	RTA22200257F.h.01.1.P.Seq	F	M00064171D:E05	CH21PRN
250	817690	RTA22200257F.e.05.1.P.Seq	F	M00064159A:H03	CH21PRN
251	819156	RTA22200256F.h.13.1.P.Seq	F	M00064106C:G03	CH21PRN
252	830904	RTA22200266F.j.12.1.P.Seq	F	M00064620D:G05	CH22PRC
253	819498	RTA22200253F.b.04.1.P.Seq	F	M00063686B:E07	CH21PRN
254	817508	RTA22200257F.g.24.1.P.Seq	F	M00064171D:E05	CH21PRN
255	817508	RTA22200252F.a.19.1.P.Seq	F	M00063515B:F06	CH21PRN
256	831160	RTA22200267F.h.01.1.P.Seq	F	M00064690A:C04	CH22PRC
257	817762	RTA22200252F.k.13.1.P.Seq	F	M00063627C:F06	CH21PRN
258	377286	RTA22200266F.k.07.1.P.Seq	F	M00064624C:B03	CH22PRC
259	831160	RTA22200267F.g.24.1.P.Seq	F	M00064690A:C04	CH22PRC
260	819994	RTA22200256F.k.11.1.P.Seq	F	M00064119C:D12	CH21PRN
261	819994	RTA22200256F.k.09.1.P.Seq	F	M00064119B:H10	CH21PRN
262	373298	RTA22200259F.c.19.1.P.Seq	F	M00063533A:C12	CH22PRC
263	819894	RTA22200256F.m.03.2.P.Seq	F	M00064126C:C02	CH21PRN
264	372718	RTA22200260F.b.22.1.P.Seq	F	M00063580D:B06	CH22PRC
265	827355	RTA22200262F.l.20.1.P.Seq	F	M00064358A:G03	CH22PRC
266	819894	RTA22200255F.d.09.1.P.Seq	F	M00063931B:E10	CH21PRN
267	827355	RTA22200266F.e.07.1.P.Seq	F	M00064601C:G07	CH22PRC
268	372718	RTA22200256F.l.03.1.P.Seq	F	M00064122C:B06	CH21PRN
269	647435	RTA22200251F.b.10.1.P.Seq	F	M00063156D:H10	CH21PRN
270	450262	RTA22200265F.a.10.1.P.Seq	F	M00064514A:G10	CH22PRC
271	484703	RTA22200255F.i.20.1.P.Seq	F	M00064032D:G04	CH21PRN
272	819498	RTA22200256F.f.12.1.P.Seq	F	M00064089B:F09	CH21PRN
273	406043	RTA22200263F.i.12.1.P.Seq	F	M00064404A:B05	CH22PRC
274	817500	RTA22200255F.f.24.1.P.Seq	F	M00063945A:C03	CH21PRN
275	818180	RTA22200264F.o.18.1.P.Seq	F	M00064506A:C07	CH22PRC
276	818143	RTA22200251F.a.03.1.P.Seq	F	M00063151A:G06	CH21PRN
277	819756	RTA22200267F.a.18.1.P.Seq	F	M00064649A:E04	CH22PRC
278	406908	RTA22200257F.i.18.3.P.Seq	F	M00064176D:H10	CH21PRN
279	124863	RTA22200256F.o.21.2.P.Seq	F	M00064136C:D12	CH21PRN
280	429009	RTA22200257F.e.24.1.P.Seq	F	M00064161B:G04	CH21PRN
281	402586	RTA22200257F.i.24.3.P.Seq	F	M00064178B:A05	CH21PRN
282	400475	RTA22200254F.i.04.1.P.Seq	F	M00063890A:H04	CH21PRN
283	403453	RTA22200264F.d.12.1.P.Seq	F	M00064450C:E07	CH22PRC
284	383021	RTA22200259F.d.06.1.P.Seq	F	M00063534C:A02	CH22PRC
285	394913	RTA22200254F.p.10.1.P.Seq	F	M00063915C:E01	CH21PRN
286	831361	RTA22200263F.k.19.1.P.Seq	F	M00064414D:D06	CH22PRC
287	646020	RTA22200267F.n.21.1.P.Seq	F	M00064723C:H04	CH22PRC

Table 1

SEQ ID	CLUSTER	SEQ NAME	ORIENT	CLONE ID	LIBRARY
288	831361	RTA22200263F.i.03.1.P.Seq	F	M00064415B:G03	CH22PRC
289	831580	RTA22200261F.f.18.1.P.Seq	F	M00064002C:H09	CH22PRC
290	402586	RTA22200257F.j.01.3.P.Seq	F	M00064178B:A05	CH21PRN
291	400475	RTA22200262F.j.21.1.P.Seq	F	M00064349D:H01	CH22PRC
292	818937	RTA22200262F.h.14.2.P.Seq	F	M00064341A:C02	CH22PRC
293	557697	RTA22200261F.j.20.1.P.Seq	F	M00064018C:E07	CH22PRC
294	831361	RTA22200265F.m.24.1.P.Seq	F	M00064569B:A09	CH22PRC
295	194490	RTA22200252F.c.10.1.P.Seq	F	M00063520D:D08	CH21PRN
296	818143	RTA22200254F.b.18.1.P.Seq	F	M00063848C:G11	CH21PRN
297	377286	RTA22200259F.a.10.1.P.Seq	F	M00063163A:G04	CH22PRC
298	831361	RTA22200265F.n.01.1.P.Seq	F	M00064569B:A09	CH22PRC
299	385307	RTA22200255F.p.07.1.P.Seq	F	M00064060B:D03	CH21PRN
300	378447	RTA22200251F.c.01.1.P.Seq	F	M00063158A:E11	CH21PRN
301	378447	RTA22200251F.b.24.1.P.Seq	F	M00063158A:E11	CH21PRN
302	817514	RTA22200260F.m.17.1.P.Seq	F	M00063968D:G08	CH22PRC
303	818942	RTA22200255F.f.03.1.P.Seq	F	M00063941B:C12	CH21PRN
304	818942	RTA22200267F.e.23.1.P.Seq	F	M00064678D:F05	CH22PRC
305	817363	RTA22200266F.f.04.1.P.Seq	F	M00064605C:G05	CH22PRC
306	818942	RTA22200255F.i.02.1.P.Seq	F	M00064025D:E07	CH21PRN
307	818942	RTA22200265F.g.23.1.P.Seq	F	M00064534D:F06	CH22PRC
308	817457	RTA22200267F.e.15.1.P.Seq	F	M00064675C:E09	CH22PRC
309	831968	RTA22200263F.f.23.1.P.Seq	F	M00064393B:H04	CH22PRC
310	530941	RTA22200253F.h.05.1.P.Seq	F	M00063785C:F03	CH21PRN
311	763446	RTA22200257F.j.05.3.P.Seq	F	M00064179A:C04	CH21PRN
312	763446	RTA22200255F.n.21.1.P.Seq	F	M00064053D:F02	CH21PRN
313	819219	RTA22200256F.f.16.1.P.Seq	F	M00064090C:A02	CH21PRN
314	763446	RTA22200258F.b.19.2.P.Seq	F	M00064248A:E02	CH21PRN
315	10154				
316	10154				

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
SEQ ID	ACCESSION	DESCRIPTION	P VALUE	ACCESSION	DESCRIPTION	P VALUE
19	<NONE>	<NONE>	<NONE>	1077580	hypothetical protein YDR125c - yeast	7
20	<NONE>	<NONE>	<NONE>	4585925	(AC007211) unknown protein	6
21	<NONE>	<NONE>	<NONE>	1085306	EVI1 protein - human	4.3
22	<NONE>	<NONE>	<NONE>	3876587	(Z81521) predicted using Genefinder; cDNA EST yk233g4.5 comes from this gene; cDNA EST yk233g4.3 comes from this gene [Caenorhabditis elegans]	0.85
23	<NONE>	<NONE>	<NONE>	1086591	(U41007) similar to S. cerevisiae nuclear protein SNF2	0.34
24	<NONE>	<NONE>	<NONE>	157272	(L11345) DNA-binding protein [Drosophila melanogaster]	0.29
25	<NONE>	<NONE>	<NONE>	2633160	(Z99108) similar to surface adhesion YfiQ [Bacillus subtilis]	0.19
26	<NONE>	<NONE>	<NONE>	755468	(U19879) transmembrane protein [Xenopus laevis]	0.042
27	<NONE>	<NONE>	<NONE>	4507339	T brachyury (mouse) homolog protein [Homo sapiens]	0.029

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
28	<NONE>	<NONE>	<NONE>	729711	PROTEASE DEGS PRECURSOR 3.4.21.-) hhoB - Escherichia coli > gi 558913 (U15661) HhoB [Escherichia coli] > gi 606174 (U18997) ORF_o355 coli] > gi 1789630 (AE000402) protease [Escherichia coli]	0.004
29	<NONE>	<NONE>	<NONE>	3168911	(AF068718) No definition line found [Caenorhabditis elegans]	8e-013
30	<NONE>	<NONE>	<NONE>	2832777	(AL021086) /prediction=(me thod:: comes from the 5' UTR [Drosophila melanogaster]	3e-040
31	X78712	H.sapiens mRNA for glycerol kinase testis specific 2	2.1	2852449	(D88207) protein kinase [Arabidopsis thaliana] > gi 2947061 (AC002521) putative protein kinase	9.1
32	X60760	L.esculentum TDR8 mRNA	2.1	157272	(L11345) DNA- binding protein [Drosophila melanogaster]	5

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
33	U40853	Oryctolagus cuniculus pulmonary surfactant protein B (SP-B) gene, complete cds	2	< NONE >	< NONE >	< NONE >
34	AF083655	Homo sapiens procollagen C-proteinase enhancer protein (PCOLCE) gene, 5' flanking region and complete cds	2	< NONE >	< NONE >	< NONE >
35	AJ223776	Staphylococcus warneri hld gene	2	< NONE >	< NONE >	< NONE >
36	U40853	Oryctolagus cuniculus pulmonary surfactant protein B (SP-B) gene, complete cds	2	< NONE >	< NONE >	< NONE >
37	X04436	Clostridium tetani gene for tetanus toxin	2	< NONE >	< NONE >	< NONE >
38	Z35787	S.cerevisiae chromosome II reading frame ORF YBL026w	2	157272	(L11345) DNA-binding protein [Drosophila melanogaster]	8.4
39	X78712	H.sapiens mRNA for glycerol kinase testis specific 2	2	2852449	(D88207) protein kinase [Arabidopsis thaliana] > gi 2947061 (AC002521) putative protein kinase	8.2

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
40	Z15056	B.subtilis genes spoVD, murE, mraY, murD	2	477124	P3A2 DNA binding protein homolog EWG - fruit fly (Drosophila melanogaster)	2.8
41	S65623	cAMP-regulated enhancer-binding protein 1 of 3]	2	119266	PROTEIN GRAINY-HEAD (DNA-BINDING PROTEIN ELF-1) (ELEMENT I-BINDING ACTIVITY) regulatory protein elf-1 - fruit fly (Drosophila melanogaster) > gi 7939 emb CAA33692 (X15657) Elf-1 protein (AA 1-1063) [Drosophila melanogaster]	0.55
42	NM_004415.1	Homo sapiens desmoplakin (DPI, DPII) (DSP) mRNA mRNA, complete cds	2	2649177	(AE001008) conserved hypothetical protein [Archaeoglobus fulgidus]	0.2
43	AF031552	Vibrio cholerae magnesium transporter (mgtE) gene, partial cds; sensor kinase (vieS), response regulator (vieA), and response regulator (vieB) genes, complete cds; and collagenase (vcc) gene,	2	2088714	(AF003139) strong similarity to NADPH oxidases; partial CDS, the gene begins in the neighboring clone	2e-013

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
		(vcc) gene, partial cds				
44	AF116852.1	Danio rerio dickkopf-1 (dkk1) mRNA, complete cds	2	3800951	(AF100657) No definition line found [Caenorhabditis elegans]	2e-019
45	X82595	P.sativum fuc gene	1.9	<NONE>	<NONE>	<NONE>
46	AF008216	Homo sapiens candidate tumor suppressor pp32r1	1.9	<NONE>	<NONE>	<NONE>
47	AF130672.1	Felis catus clone Fca603 microsatellite sequence	1.9	<NONE>	<NONE>	<NONE>
48	AJ007044	Oryctolagus Cuniculus sod gene	1.9	388055	(L22981) merozoite surface protein-1 [Plasmodium chabaudi]	7.8
49	AC004497	Homo sapiens chromosome 21, P1 clone LBNL#6	1.9	160925	(M94346) A.1.12/9 antigen [Schistosoma mansoni]	7.7

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
50	U30290	Rattus norvegicus galanin receptor GALR1 mRNA, complete cds	1.9	3024079	GALECTIN-4 (LACTOSE-BINDING LECTIN 4) (L-36 LACTOSE BINDING PROTEIN) (L36LBP) > gi 2281707 sapiens] > gi 2623387 (U82953) galectin-4 [Homo sapiens]	4.5
51	Y13234	Chironomus tentans mRNA for chitinase, 1695 bp	1.9	4567068	(AF125568) tumor suppressing STF cDNA 4 [Homo sapiens]	3.4
52	NM_003644.1	Homo sapiens growth arrest-specific 7 (GAS7) mRNA > :: emb AJ224876 HSAJ4876 Homo sapience mRNA for GAS7 protein	1.9	125560	PROTEIN KINASE C, GAMMA TYPE C (EC 2.7.1.-) gamma - rabbit > gi 165652 (M19338) protein kinase delta [Oryctolagus cuniculus]	0.53
53	AB013448.1	Oryza sativa gene for Pib, complete cds	1.8	<NONE>	<NONE>	<NONE> >
54	D63854	Human cytomegalovirus DNA, replication origin	1.8	<NONE>	<NONE>	<NONE> >
55	AB002340	Human mRNA for KIAA0342 gene, complete cds	1.8	<NONE>	<NONE>	<NONE> >
56	AF017779	Mus musculus vitamin D receptor gene, promoter region	1.8	<NONE>	<NONE>	<NONE> >

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
57	D63854	Human cytomegalovirus DNA, replication origin	1.8	<NONE>	<NONE>	<NONE>
58	M24102	Bovine ADP/ATP translocase T1 mRNA, complete cds.	1.8	<NONE>	<NONE>	<NONE>
59	AC004497	Homo sapiens chromosome 21, P1 clone LBNL#6	1.8	<NONE>	<NONE>	<NONE>
60	M37394	Rat epidermal growth factor receptor mRNA.	1.8	<NONE>	<NONE>	<NONE>
61	AF006304	Saccharomyces cerevisiae protein tyrosine phosphatase (PTP3) gene, complete cds	1.8	<NONE>	<NONE>	<NONE>
62	D13454	Candida albicans CACHS3 gene for chitin synthase III	1.8	<NONE>	<NONE>	<NONE>
63	Y00354	Xenopus laevis gene encoding vitellogenin A2	1.8	1077580	hypothetical protein YDR125c - yeast	7.5
64	U90936	Aspergillus niger px27 gene, promoter region	1.8	4337033	(AF124138) transcriptional activator protein CdaR [Streptomyces coelicolor] transcriptional regulator [Streptomyces coelicolor]	7.3

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
65	D84448	Cavia cobaya mRNA for Na ⁺ ,K ⁺ -ATPase beta-3 subunit, complete cds	1.8	4704603	(AF109916) putative dehydrin	7.1
66	AF039948	Xenopus laevis clone H-0 transcription elongation factor S-II (TFIIS) precursor RNA, isoform TFIIS.h, partial cds	1.8	1695839	(U58151) envelope glycoprotein [Human immunodeficiency virus type 1]	5.6
67	M18061	Xenopus laevis vitellogenin gene, complete cds.	1.8	780502	(U18466) AP endonuclease class II [African swine fever virus] >gi 1097525 p rf 2113434ET AP endonuclease:IS OTYPE=class II [African swine fever virus]	3.1
68	U61112	Mus musculus Eya3 homolog mRNA, complete cds	1.8	3043646	(AB011133) KIAA0561 protein [Homo sapiens]	1.9
69	AB018442	Oryza sativa mRNA for phytochrome C, complete cds	1.8	4455041	(AF116463) unknown [Streptomyces lincolnensis]	0.49

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
70	D63854	Human cytomegalovirus DNA, replication origin	1.8	1169200	DNA-DAMAGE-REPAIR/TOLE RATION PROTEIN DRT111 PRECURSOR > gi 421829 pi r S33706 DNA-damage resistance protein - Arabidopsis thaliana and DNA-damage resistance protein (DRT111) mRNA, complete cds.], gene product [Arabidopsis thaliana]	0.22
71	D26549	Bovine mRNA for adseverin, complete cds	1.8	755468	(U19879) transmembrane protein [Xenopus laevis]	0.042
72	J05211	Human desmoplakin mRNA, 3' end.	1.8	728867	ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR > gi 99694 pir S21961 proline-rich protein APG - Arabidopsis thaliana > gi 22599 em b CAA42925	0.015

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
73	NM_004415 .1	Homo sapiens desmoplakin (DPI, DPII) (DSP) mRNA, complete cds	1.8	728867	ANTER-SPECIFIC PROLINE-RICH PROTEIN APG PRECURSOR > gi 99694 pir S21961 proline-rich protein APG - Arabidopsis thaliana > gi 22599 emb b CAA42925	0.015
74	AF038604	Caenorhabditis elegans cosmid B0546	1.8	3877951	(Z81555) predicted using Genefinder	3e-008
75	AF038604	Caenorhabditis elegans cosmid B0546	1.8	3877951	(Z81555) predicted using Genefinder	2e-011
76	U23551	Prochlorothrix hollandica phosphomannomutase	1.8	2828280	(AL021687) putative protein [Arabidopsis thaliana] > gi 2832633 emb b CAA16762 (AL021711) putative protein [Arabidopsis thaliana]	2e-013
77	S60150	ORF1...ORF6 {3' terminal reigon} [chrysanthemum virus B CVB, Genomic RNA, 6 genes, 3426 nt]	1.8	1065454	(U40410) C54G7.2 gene product [Caenorhabditis elegans]	2e-019
78	AB014558	Homo sapiens mRNA for KIAA0658 protein, partial cds	1.8	3850072	(AL033385) dna-directed rna polymerase iii subunit [Schizosaccharomyces pombe]	6e-027

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
79	X17191	E.gracilis chloroplast RNA polymerase rpoB-rpoC1-rpoC2 operon	1.7	<NONE>	<NONE>	<NONE>
80	X07729	R.norvegicus gene encoding neuron-specific enolase, exons 8-12	1.7	4584544	(AL049608) extensin-like protein	8.8
81	D38178	Human gene for cytosolic phospholipase A2, exon 1	1.7	73714	infected cell protein ICP34.5 - human herpesvirus 1 (strain F) > gi 330123 (M12240) infected cell protein [Herpes simplex virus type 1]	1.1
82	U23551	Prochlorothrix hollandica phosphomannomutase	1.7	2828280	(AL021687) putative protein [Arabidopsis thaliana] > gi 2832633 emb CAA16762 (AL021711) putative protein [Arabidopsis thaliana]	2e-010
83	Y00525	Klebsiella pneumoniae nifL gene for regulatory protein	1.6	3800951	(AF100657) No definition line found [Caenorhabditis elegans]	6e-013
84	AF100170.1	Bos taurus major fibrous sheath protein precursor, mRNA, complete cds	1.5	463552	(U05877) AF-1 [Homo sapiens]	0.074
85	Y13441	Homo sapiens Rox gene, exon 2	0.74	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
86	L46792	Actinidia deliciosa clone AdXET-5 xyloglucan endotransglycosylase precursor (XET) mRNA, complete cds	0.73	3170252	(AF043636) circumsporozoite protein [Plasmodium chabaudi]	0.001
87	U73489	Drosophila melanogaster Nem (nem) mRNA, complete cds	0.7	3915994	HYPOTHETICAL 53.2 KD PROTEIN IN PRC-PRPA INTERGENIC REGION	3e-005
88	U95097	Xenopus laevis mitotic phosphoprotein 43 mRNA, partial cds	0.68	157272	(L11345) DNA-binding protein [Drosophila melanogaster]	8.5
89	AF082012	Caenorhabditis elegans UDP-N-acetylglucosamine:alpha-3-D-mannoside b-1,2-N-acetylglucosaminyltransferase I (gly-14) mRNA, complete cds	0.67	2494313	PUTATIVE TRANSLATION INITIATION FACTOR EIF-2B SUBUNIT 1 (EIF-2B GDP-GTP EXCHANGE FACTOR) eIF-2B, subunit alpha - Methanococcus jannaschii aIF-2B, subunit delta (aIF2BD) [Methanococcus jannaschii]	8.4
90	U04354	Mus musculus ADSEVERIN mRNA, complete cds	0.67	4755188	(AC007018) unknown protein	8e-026

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
91	M68881	S.pombe cigl + gene, complete cds.	0.67	2078441	(U56964) weak similarity to S. cerevisiae intracellular protein transport protein US)1 (SP:P25386)	2e-030
92	U95097	Xenopus laevis mitotic phosphoprotein 43 mRNA, partial cds	0.66	2829685	PROTEIN-TYROSINE PHOSPHATASE X PRECURSOR (R-PTP-X) (PTP 1A-2BETA) (PROTEIN TYROSINE PHOSPHATASE NP) (PTP-NP) > gi 1515425 (U57345) protein tyrosine phosphatase-NP [Mus musculus]	6.2
93	Z15056	B.subtilis genes spoVD, murE, mraY, murD	0.66	477124	P3A2 DNA binding protein homolog EWG - fruit fly (Drosophila melanogaster)	2.1
94	M86808	Human pyruvate dehydrogenase complex (PDHA2) gene, complete cds.	0.65	< NONE >	< NONE >	< NONE >
95	J03754	Rat plasma membrane Ca2+ ATPase-isoform 2 mRNA, complete cds.	0.65	4507549	transmembrane protein with EGF-like and two follistatin-like domains 1 > gi 755466	8e-006

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
96	NM_000887.1	Homo sapiens integrin, alpha X (antigen CD11C emb Y00093 H SP15095 H.sapiens mRNA for leukocyte adhesion glycoprotein p150,95	0.64	<NONE>	<NONE>	<NONE>
97	L27080	Human melanocortin 5 receptor (MC5R) gene, complete cds.	0.64	<NONE>	<NONE>	<NONE>
98	U07890	Mus musculus C57BL/6J epidermal surface antigen (mesa) mRNA, complete cds.	0.64	<NONE>	<NONE>	<NONE>
99	AF079139	Streptomyces venezuelae pikCD operon, complete sequence	0.64	3041869	(U96109) proline-rich transcription factor ALX3 [Mus musculus]	2.8
100	M16140	Chicken ovoinhibitor gene, exon 15.	0.64	123984	ACROSIN INHIBITORS IIA AND IIB	4e-008
101	NM_000887.1	Homo sapiens integrin, alpha X (antigen CD11C emb Y00093 H SP15095 H.sapiens mRNA for leukocyte adhesion glycoprotein p150,95	0.63	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
102	Z17316	Kluyveromyces lactis for gene encoding phosphofructokinase beta subunit	0.63	< NONE >	< NONE >	< NONE >
103	Z25470	H.sapiens melanocortin 5 receptor gene, complete CDS	0.63	< NONE >	< NONE >	< NONE >
104	L19954	Bacillus subtilis feuA, B, and C genes, 3 ORFs, 2 complete cds's and 5'end.	0.63	< NONE >	< NONE >	< NONE >
105	U44405	Spiroplasma citri chromosome pre-inversion border, SPV1-like sequences, transposase gene, partial cds, adhesin-like protein P58 gene, complete cds.	0.63	2499642	SERINE/THREONINE-PROTEIN KINASE STE20 HOMOLOG > gi 1737181 (U73457) Cst20p [Candida albicans]	7.7

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
106	Z28264	S.cerevisiae chromosome XI reading frame ORF YKR039w	0.63	3880930	(AL021481) similar to Phosphoglucomutase and phosphomannomutase phosphoserine; cDNA EST EMBL:D36168 comes from this gene; cDNA EST EMBL:D70697 comes from this gene; cDNA EST yk373h9.5 comes from this gene; cDNA EST EMBL:T00805 ...	2e-014
107	AE001107	Archaeoglobus fulgidus section 172 of 172 of the complete genome	0.62	<NONE>	<NONE>	<NONE>
108	Z14112	B.firmus TopA gene encoding DNA topoisomerase I	0.62	310115	(L02530) Drosophila polarity gene (frizzled) homologue	0.026
109	AF118101	Toxoplasma gondii protein kinase 6 (tpk6) mRNA, complete cds	0.62	726403	(U23175) similar to anion exchange protein [Caenorhabditis elegans]	4e-018
110	M59743	Rabbit cardiac muscle Ca-2+ release channel	0.61	<NONE>	<NONE>	<NONE>
111	M12036	Human tyrosine kinase-type receptor (HER2) gene, partial cds.	0.61	61962	(X58484) gag [Simian foamy virus]	7.5

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
112	AF043195	Homo sapiens tight junction protein ZO (ZO-2) gene, alternative splice products, promoter and exon A	0.61	1572629	(U69699) unknown protein precursor [Mus musculus]	7.5
113	U18178	Human HLA class I genomic survey sequence.	0.61	1336688	(S81116) properdin [guinea pigs, spleen, Peptide, 470 aa] [Cavia]	5.7
114	U44405	Spiroplasma citri chromosome pre-inversion border, SPV1-like sequences, transposase gene, partial cds, adhesin-like protein P58 gene, complete cds.	0.61	2827531	(AL021633) hypothetical protein	3.3
115	Z33011	M.capricolum DNA for CONTIG MC008	0.61	3915729	HYPERPLASTIC DISCS PROTEIN (HYDROPHILIC PROTEIN) > gi 2673887 (L14644) hyperplastic discs protein	0.26
116	NM_001429.1	Homo sapiens E1A binding protein p300 mRNA, complete cds. > :: gb I62297 I62297 Sequence 1 from patent US 5658784	0.61	4204294	(AC003027) lcl prt_seq No definition line found	5e-005

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
117	Z25418	C.familiaris MHC class Ib gene (DLA-79) gene, complete CDS	0.61	3877493	(Z48583) similar to ATPases associated with various cellular activities (AAA); cDNA EST EMBL:Z14623 comes from this gene; cDNA EST EMBL:D75090 comes from this gene; cDNA EST EMBL:D72255 comes from this gene; cDNA EST yk200e4.5 ...	1e-007
118	AB002150	Bacillus subtilis DNA for FeuB, FeuA, YbbB, YbbC, YbbD, YbzA, YbbE, YbbF, YbbH, YbbI, YbbJ, YbbK, YbbL, YbbM, YbbP, complete cds	0.6	< NONE >	< NONE >	< NONE >
119	Y07786	V.cholerae ORF's involved in lipopolysaccharide synthase	0.6	< NONE >	< NONE >	< NONE >
120	Z17316	Kluyveromyces lactis for gene encoding phosphofructokinase beta subunit	0.6	< NONE >	< NONE >	< NONE >

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
121	Z71403	S.cerevisiae chromosome XIV reading frame ORF YNL127w	0.6	<NONE>	<NONE>	<NONE>
122	L34641	Homo sapiens platelet/endothelial cell adhesion molecule-1 (PECAM-1) gene, exon 10.	0.6	1147634	(U42213) micronemal TRAP-C1 protein homolog	9.6
123	AF070572	Homo sapiens clone 24778 unknown mRNA	0.6	399034	N-ACETYLMURAMOYL-L-ALANINE AMIDASE AMIB PRECURSOR > gi 628763 pir S41741 N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28) - Escherichia coli > gi 304914 (L19346) N-acetylmuramoyl-L-alanine amidase [Escherichia coli] N-acetylmuramoyl-L-alanine amidase II; a	2.5
124	X75627	C.burnetii trxB, spoIIIE and serS genes	0.6	3036833	(AJ003163) apsB [Emericella nidulans]	0.28
125	Z99765	Flaveria pringlei gdcsh gene	0.59	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
126	U02538	Mycoplasma hyopneumoniae J ATCC 25934 23S rRNA gene, partial sequence	0.59	<NONE>	<NONE>	<NONE>
127	Z71403	S.cerevisiae chromosome XIV reading frame ORF YNL127w	0.59	<NONE>	<NONE>	<NONE>
128	X03942	Mouse simple repetitive DNA (sqr family) transcript (clone pmlc 2) with conserved GACA/GATA repeats	0.59	<NONE>	<NONE>	<NONE>
129	U11844	Mus musculus glucose transporter (GLUT3) gene, exon 1	0.59	<NONE>	<NONE>	<NONE>
130	D63395	Homo sapiens mRNA for NOTCH4, partial cds	0.59	4433616	(AF107018) alpha-mannosidase IIx [Mus musculus]	1.8
131	Z33011	M.capricolum DNA for CONTIG MC008	0.59	3915729	HYPERPLASTIC DISCS PROTEIN (HYD PROTEIN) > gi 2673887 (L14644) hyperplastic discs protein	0.27
132	U05670	Haemophilus influenzae DL42 Lex2A and Lex2B genes, complete cds.	0.58	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
133	L27080	Human melanocortin 5 receptor (MC5R) gene, complete cds.	0.58	123984	ACROSIN INHIBITORS IIA AND IIB	2e-006
134	AF043195	Homo sapiens tight junction protein ZO (ZO-2) gene, alternative splice products, promoter and exon A	0.57	1572629	(U69699) unknown protein precursor [Mus musculus]	6.7
135	U57707	Bos taurus activin receptor type IIB precursor	0.57	807646	(M17294) unknown protein [Human herpesvirus 4]	0.068
136	Z17316	Kluyveromyces lactis for gene encoding phosphofructokinase beta subunit	0.56	<NONE>	<NONE>	<NONE>
137	M21535	Human erg protein (ets-related gene) mRNA, complete cds.	0.56	<NONE>	<NONE>	<NONE>
138	M64932	Candida maltosa cyclohexamide resistance protein	0.56	3219524	(AF069428) NADH dehydrogenase subunit IV [Alligator mississippiensis] > gi 3367630 emb CAA73570 (Y13113) NADH dehydrogenase subunit 4 [Alligator mississippiensis]	1.3

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
139	AE000342	Escherichia coli K-12 MG1655 section 232 of 400 of the complete genome	0.56	3874685	(Z78539) Similarity to S.pombe hypothetical protein C4G8.04 (SW:YAD4_SC HPO); cDNA EST EMBL:D27846 comes from this gene; cDNA EST EMBL:D27845 comes from this gene; cDNA EST yk202h7.3 comes from this gene; cDNA EST yk202h7.5 come...	0.088
140	Z15056	B.subtilis genes spoVD, murE, mraY, murD	0.55	477124	P3A2 DNA binding protein homolog EWG - fruit fly (Drosophila melanogaster)	3.7
141	Z58167	H.sapiens CpG island DNA genomic MseI fragment, clone 30e10, forward read cpg30e10.ft1b	0.53	<NONE>	<NONE>	<NONE>
142	M27159	Rat potassium channel-Kv2 gene, partial cds.	0.53	1850920	(U21247) Bet [Human spumaretrovirus]	0.9
143	M15555	Mouse Ig germline V-kappa-24 chain (VK24C) gene, exons 1 and 2.	0.24	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
144	U95097	Xenopus laevis mitotic phosphoprotein 43 mRNA, partial cds	0.24	399109	TRANSCRIPTI ON FACTOR BF-1 (BRAIN FACTOR 1) (BF1) > gi 92020 pir JH0672 brain factor 1 protein - rat > gi 203135 (M87634) BF-1 [Rattus norvegicus]	4
145	AJ002014	Crythecodinium cohnii mRNA for nuclear protein JUS1	0.24	416704	BALBIANI RING PROTEIN 3 PRECURSOR balbiani ring 3 (BR3) [Chironomus tentans]	0.36
146	L35330	Rattus norvegicus glutathione S-transferase Yb3 subunit gene, complete cds.	0.23	1388158	(U58204) myomesin [Gallus gallus]	8.8
147	NM_001432.1	Homo sapiens epiregulin (EREG) mRNA > :: dbj D30783 D30783 Homo sapiens mRNA for epiregulin, complete cds	0.23	2851520	TRANSFORMI NG GROWTH FACTOR ALPHA PRECURSOR (TGF-ALPHA) (EGF-LIKE TGF) (ETGF) (TGF TYPE 1) precursor - rat > gi 207282 (M31076) transforming growth factor alpha precursor [Rattus norvegicus]	2e-008

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
148	U57043	Cebus apella gamma globin (gamma1) gene, complete cds	0.22	<NONE>	<NONE>	<NONE>
149	AB023188.1	Homo sapiens mRNA for KIAA0971 protein, complete cds	0.22	<NONE>	<NONE>	<NONE>
150	M18105	Yeast (S.cerevisiae) SST2 gene encoding desensitization to alpha- factor pheromone, complete cds.	0.22	<NONE>	<NONE>	<NONE>
151	AJ001113	Homo sapiens UBE3A gene, exon 16	0.22	3122961	ENHANCER OF SPLIT GROUCHO-LIKE PROTEIN 1 > gi 2408145 (U18775) enhancer of split groucho	8.5
152	L35330	Rattus norvegicus glutathione S-transferase Yb3 subunit gene, complete cds.	0.22	1388158	(U58204) myomesin [Gallus gallus]	8.1
153	D42042	Human mRNA for KIAA0085 gene, partial cds	0.22	4827063	zinc finger protein 142 (clone pHZ-49) > gi 3123312 s p P52746 Z142 HUMAN ZINC FINGER PROTEIN 142 (KIAA0236) (HA4654) > gi 1510147 d bj BAA13242	6.1

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
154	L35330	Rattus norvegicus glutathione S-transferase Yb3 subunit gene, complete cds.	0.22	2853301	(AF007194) mucin [Homo sapiens]	1.6
155	Z11653	H.sapiens DBH gene complex repeat polymorphism DNA	0.22	3819705	(AL032824) syntaxin binding protein 1; sec1 family secretory protein [Schizosaccharomyces pombe]	1.2
156	L29063	Candida albicans fatty acid synthase alpha subunit (FAS2) gene, complete cds.	0.22	3046871	(AB003753) high sulfur protein B2E [Rattus norvegicus]	0.32
157	M64865	Horse alcohol dehydrogenase-S-isoenzyme mRNA, complete cds.	0.22	2213909	(AF004874) latent TGF-beta binding protein-2 [Mus musculus]	0.037
158	Y09472	B.taurus gene encoding preprododecapeptide	0.21	2909874	(AF047829) melatonin-related receptor [Ovis aries]	7.6
159	Y09472	B.taurus gene encoding preprododecapeptide	0.21	2909874	(AF047829) melatonin-related receptor [Ovis aries]	7.5
160	X80301	N.tabacum axi 1 gene	0.21	2832715	(AJ003066) subunit beta of the mitochondrial fatty acid beta-oxidation multienzyme complex [Bos taurus]	6

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
161	AF073485	Homo sapiens MHC class I-related protein MR1 precursor (MR1) gene, partial cds	0.21	2224559	(AB002307) KIAA0309 [Homo sapiens]	3.3
162	S78251	growth hormone receptor {alternatively spliced, exon 1B} [sheep, Merino, skeletal muscle, mRNA. Partial, 438 nt]	0.21	729381	DYNAMIN-1 (DYNAMIN BREDNM19)	2
163	U16135	Synechococcus sp. Clp protease proteolytic subunit	0.21	135514	T-CELL RECEPTOR BETA CHAIN PRECURSOR precursor (ANA 11) - rabbit	0.02
164	X95601	M.hominis Imp3 and Imp4 genes	0.21	2995445	(Y10496) CDV-1 protein [Mus musculus]	0.005
165	X95601	M.hominis Imp3 and Imp4 genes	0.21	2995447	(Y10495) CDV-1R protein [Mus musculus]	0.005
166	AF124249.1	Homo sapiens SH2-containing protein Nsp1 mRNA, complete cds	0.21	423456	epidermal growth factor-receptor-binding protein GRB-4 - mouse (fragment)	8e-010
167	AF030282	Danio rerio homeobox protein Six7 (six7) mRNA, complete cds	0.21	3928083	(AC005770) unknown protein [Arabidopsis thaliana]	2e-014
168	X83427	O.anatinus mitochondrial DNA, complete genome	0.21	132575	RIBONUCLEASE INHIBITOR	3e-021

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
169	AJ001113	Homo sapiens UBE3A gene, exon 16	0.2	<NONE>	<NONE>	<NONE>
170	AF081533.1	Anopheles gambiae putative gram negative bacteria binding protein gene, complete cds	0.2	<NONE>	<NONE>	<NONE>
171	U70316	Dictyostelium discoideum Iona (iona) gene, partial cds	0.2	<NONE>	<NONE>	<NONE>
172	AF009341	Homo sapiens E6-AP ubiquitin-protein ligase	0.2	<NONE>	<NONE>	<NONE>
173	L35330	Rattus norvegicus glutathione S-transferase Yb3 subunit gene, complete cds.	0.2	3702275	(AC005793) KIAA0561 protein [AA 1-593] [Homo sapiens]	2.5
174	AE000573.1	Helicobacter pylori 26695 section 51 of 134 of the complete genome	0.2	3947855	(AL034381) putative Golgi membrane protein	2.5
175	X83230	G.gallus hsp90beta gene	0.2	3258596	(U95821) putative transmembrane GTPase [Drosophila melanogaster]	0.81
176	X57157	Chicken mRNA for Hsp47, heat shock protein 47	0.2	108325	insulin-like growth factor-binding protein 6	0.17

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
177	M58748	Chicken alpha-globin gene domain with structural matrix attachment sites.	0.2	1086863	(U41272) T03G11.6 gene product [Caenorhabditis elegans]	4e-005
178	AB016815	Anthocidaris crassispina mRNA for Src-type protein tyrosine kinase, complete cds	0.2	423456	epidermal growth factor-receptor-binding protein GRB-4 - mouse (fragment)	1e-012
179	AF030282	Danio rerio homeobox protein Six7 (six7) mRNA, complete cds	0.2	3928083	(AC005770) unknown protein [Arabidopsis thaliana]	3e-014
180	AL035559	Streptomyces coelicolor cosmid 9F2	0.2	2088714	(AF003139) strong similarity to NADPH oxidases; partial CDS, the gene begins in the neighboring clone	3e-022
181	S79641	SDH=succinate dehydrogenase flavoprotein subunit Mutant, 387 nt]	0.2	4755188	(AC007018) unknown protein	2e-022
182	X75383	H.sapiens mRNA for TFIIA-alpha	0.19	< NONE >	< NONE >	< NONE >
183	U53901	Hippopotamus amphibius b-casein gene, exon 7, partial cds	0.19	< NONE >	< NONE >	< NONE >
184	J05265	Mouse interferon gamma receptor mRNA, complete cds.	0.19	77356	hypothetical 70K protein - eggplant mosaic virus	0.0005

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
185	U72353	Rattus norvegicus lamin B1 mRNA, complete cds	0.19	3880857	(AL031633) cDNA EST yk404d1.5 comes from this gene; cDNA EST yk404d1.3 comes from this gene	2e-006
186	AB016815	Anthocidaris crassispina mRNA for Src-type protein tyrosine kinase, complete cds	0.19	3930217	(AF047487) Nck-2 [Homo sapiens]	2e-007
187	D10911	Mus musculus DNA for MS2 protein, complete cds	0.19	2662366	(D86332) membrane type-2 matrix metalloproteinase [Mus musculus]	5e-011
188	AB015345	Homo sapiens HRIHFB2216 mRNA, partial cds	0.075	3877417	(Z66564) similar to anion exchange protein	6.4
189	AF086410	Homo sapiens full length insert cDNA clone ZD77B03	0.075	3023371	PEROMONE B BETA 1 RECEPTOR	4.9
190	K02024	Human T-cell lymphotropic virus type II env gene encoding envelope glycoprotein, complete cds.	0.075	2791527	(AL021246) PE_PGRS [Mycobacterium tuberculosis]	0.11

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
191	M10188	X.laevis mitochondrial DNA containing the D-loop, and the 12S rRNA, apocytochrome b, Glu-tRNA, Thr-tRNA, Pro-tRNA and Phe-tRNA genes.	0.074	4753163	huntingtin DISEASE PROTEIN) (HD PROTEIN) > gi 454415 (L12392) Huntington's Disease protein [Homo sapiens]	2.8
192	X85525	G.gallus AG repeat region (GgaMU130)	0.073	984339	(U20966) Rev [Simian immunodeficiency virus]	3.6
193	AJ238394.1	Homo sapiens AML2 gene (partial)	0.07	4240219	(AB020672) KIAA0865 protein [Homo sapiens]	2
194	AF039704	Homo sapiens lysosomal pepstatin insensitive protease (CLN2) gene, complete cds	0.069	2894106	(Z78279) Collagen alpha1 [Rattus norvegicus]	0.39
195	K02024	Human T-cell lymphotropic virus type II env gene encoding envelope glycoprotein, complete cds.	0.068	4504857	potassium intermediate/small conductance calcium-activated channel, subfamily N, member 3 > gi 3309531 (AF031815) calcium-activated potassium channel [Homo sapiens]	0.5

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
196	Z60719	H.sapiens CpG island DNA genomic MseI fragment, clone 33a11, forward read cpg33a11.ft1m	0.068	4826874	nucleoporin 214kD (CAIN) PROTEIN NUP214 (NUCLEOPORIN NUP214) (214 KD NUCLEOPORIN) transforming protein (can) - human sapiens]	0.044
197	AF053994	Lycopersicon esculentum Hcr2-0A (Hcr2-0A) gene, complete cds	0.068	2842699	PUTATIVE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE C6G9.08 (UBIQUITIN THIOLESTERASE) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE)	9e-009
198	AJ233650.1	Equus caballus endogenous retroviral sequence ERV-L pol gene, clone ERV-L Horse1	0.067	< NONE >	< NONE >	< NONE >
199	M10188	X.laevis mitochondrial DNA containing the D-loop, and the 12S rRNA, apocytochrome b, Glu-tRNA, Thr-tRNA, Pro-tRNA and Phe-tRNA genes.	0.067	4753163	huntingtin DISEASE PROTEIN) (HD PROTEIN) > gi 454415 (L12392) Huntington's Disease protein [Homo sapiens]	2.5

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
200	U14646	Murine hepatitis virus Y strain S glycoprotein gene, complete cds.	0.067	3880930	(AL021481) similar to Phosphoglucosyltransferase and phosphomannosyltransferase phosphoserine; cDNA EST EMBL:D36168 comes from this gene; cDNA EST EMBL:D70697 comes from this gene; cDNA EST yk373h9.5 comes from this gene; cDNA EST EMBL:T00805 ...	1e-019
201	X15373	Mouse cerebellum mRNA for P400 protein	0.066	164507	(M81771) immunoglobulin gamma-chain [Sus scrofa]	9.4
202	AF086410	Homo sapiens full length insert cDNA clone ZD77B03	0.066	3023371	PEROMYCEIN RECEPTOR	4.2
203	AL034492	Streptomyces coelicolor cosmid 6C5	0.066	3800951	(AF100657) No definition line found [Caenorhabditis elegans]	3e-015
204	L13377	Staphylococcus aureus enterotoxin gene, 3' end.	0.065	<NONE>	<NONE>	<NONE>
205	U83478	Thelephoraceae sp. 'Taylor #13' ITS1, 5.8S ribosomal RNA gene, and ITS2, complete sequence	0.065	3877335	(Z92786) predicted using Genefinder	9.1

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
206	AJ002014	Crythecodinium cohnii mRNA for nuclear protein JUS1	0.065	1213283	(U40576) SIM2 [Mus musculus]	0.47
207	AB016804	Aloe arborescens mRNA for NADP-malic enzyme, complete cds	0.065	2832777	(AL021086) /prediction=(method:; comes from the 5' UTR [Drosophila melanogaster]	5e-036
208	AJ002014	Crythecodinium cohnii mRNA for nuclear protein JUS1	0.063	1213283	(U40576) SIM2 [Mus musculus]	0.45
209	AB023143.1	Homo sapiens mRNA for KIAA0926 protein, complete cds	0.024	132575	RIBONUCLEASE INHIBITOR	8e-026
210	U72966	Human hepatocyte nuclear factor 4-alpha gene, exon 7	0.022	<NONE>	<NONE>	<NONE>
211	X02801	Mouse gene for glial fibrillary acidic protein	0.022	2231607	(U85917) nef protein [Human immunodeficiency virus type 1]	7
212	AF017636	Mesocricetus auratus 3-ketosteroid reductase	0.022	2723362	(AF023459) lustrin A [Haliotis rufescens]	0.097
213	Z36879	F.pringlei gdcSPA gene for P-protein of the glycine cleavage system	0.008	<NONE>	<NONE>	<NONE>
214	X73150	P.sativum GapC1 gene	0.008	1572629	(U69699) unknown protein precursor [Mus musculus]	8.6

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
215	AJ239031.1	Homo sapiens LSS gene, partial, exons 22, 23 and joined CDS	0.008	4508019	zinc finger protein 231 protein [Homo sapiens]	0.01
216	U76602	Human 180 kDa bullous pemphigoid antigen 2/type XVII collagen (BPAG2/COL17A1) gene, exons 49, 50, 51 and 52	0.007	3170252	(AF043636) circumsporozoite protein [Plasmodium chabaudi]	0.0001
217	M11283	Aplysia californica FMRFamide mRNA, partial cds, clone FMRF-2.	0.007	3874685	(Z78539) Similarity to S.pombe hypothetical protein C4G8.04 (SW:YAD4_SC HPO); cDNA EST EMBL:D27846 comes from this gene; cDNA EST EMBL:D27845 comes from this gene; cDNA EST yk202h7.3 comes from this gene; cDNA EST yk202h7.5 come...	9e-013
218	J03998	P.falciparum glutamic acid-rich protein gnen, complete cds.	0.003	<NONE>	<NONE>	<NONE>
219	Z23143	M.musculus ALK-6 mRNA, complete CDS	0.002	2393890	(AF006064) protein kinase homolog [Fowlpox virus]	1e-011

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
220	AB007914	Homo sapiens mRNA for KIAA0445 protein, complete cds	0.001	2136964	cysteine-rich hair keratin associated protein - rabbit > gi 510541 emb CAA56339 (X80035) cysteine rich hair keratin associated protein	1.9
221	AB012105	Brassica rapa mRNA for SLG45, complete cds	0.0008	3687246	(AC005169) putative suppressor protein [Arabidopsis thaliana]	5.5
222	L41608	Methylobacterium extorquens (clone pDN9, HINDIIIAB) mxaS gene 3' end, mxaA, mxaC, mxaK, mxaL and mxaD genes, complete cds.	0.0008	3024235	NERVOUS-SYSTEM SPECIFIC OCTAMER-BINDING TRANSCRIPTION FACTOR N-OCT 3 PROTEIN)	5.1
223	AB007914	Homo sapiens mRNA for KIAA0445 protein, complete cds	0.0008	2136964	cysteine-rich hair keratin associated protein - rabbit > gi 510541 emb CAA56339 (X80035) cysteine rich hair keratin associated protein	2.5
224	AC002293	Genomic sequence from Human 9q34, complete sequence [Homo sapiens]	0.0008	2789557	(AF034316) MHC class I antigen [Triakis scyllium]	0.0002

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
225	L16013	Rattus norvegicus Q-like gene sequence	9e-005	< NONE >	< NONE >	< NONE >
226	AF148512.1	Homo sapiens hexokinase II gene, promoter region	9e-005	< NONE >	< NONE >	< NONE >
227	U94776	Human muscle glycogen phosphorylase (PYGM) gene, exons 6 through 17	9e-005	4759138	solute carrier family 7 transporter 3 [Homo sapiens]	5.4
228	X56030	H.sapiens IAPP gene for amyloid polypeptide, exon 1	1e-005	< NONE >	< NONE >	< NONE >
229	U36515	Human CT microsatellite, clone GM5927-CT-2-3, from the tandemly repeated genes encoding U2 small nuclear RNA (RNU2 locus)	4e-007	2435616	(AF026215) No definition line found [Caenorhabditis elegans]	0.85
230	AB011119	Homo sapiens mRNA for KIAA0547 protein, complete cds	4e-007	4758508	airway trypsin-like protease protease [Homo sapiens]	3e-031
231	NM_000521.1	Homo sapiens hexosaminidase B (beta polypeptide) (HEXB) mRNA	5e-008	2119379	slow muscle troponin T - chicken T [Gallus gallus]	2.8
232	X13895	Human serum amyloid A (GSAA1) gene, complete cds	4e-008	699405	(U18682) novel antigen receptor [Ginglymostoma cirratum]	7.7

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
233	AB009288.1	Homo sapiens mRNA for N-copine, complete cds	4e-008	4520342	(AB008893) N-copine [Mus musculus]	3e-006
234	AB011119	Homo sapiens mRNA for KIAA0547 protein, complete cds	4e-008	4758508	airway trypsin-like protease [Homo sapiens]	1e-028
235	X13895	Human serum amyloid A (GSAA1) gene, complete cds	5e-009	699405	(U18682) novel antigen receptor [Ginglymostoma cirratum]	7.8
236	X13895	Human serum amyloid A (GSAA1) gene, complete cds	2e-009	699405	(U18682) novel antigen receptor [Ginglymostoma cirratum]	7.2
237	U64997	Bos taurus ribonuclease K6 gene, partial cds	2e-009	3914810	RIBONUCLEASE K6 PRECURSOR (RNASE K6) > gi 2745760 (AF037086) ribonuclease k6 precursor	3e-018
238	J02635	Rat liver alpha-2-macroglobulin mRNA, complete cds.	2e-009	112913	ALPHA-2-MACROGLOBULIN PRECURSOR precursor - rat > gi 202592 (J02635) prealpha-2-macroglobulin [Rattus norvegicus]	4e-019
239	Z78141	M.musculus partial cochlear mRNA (clone 29C9)	5e-010	3219569	(AL023893) /prediction=(method::;	4e-009

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
240	AF060917	Gambusia affinis microsatellite Gafu6	2e-010	3874618	(Z48241) similar to coiled coil domains; cDNA EST yk302g12.5 comes from this gene; cDNA EST yk365d10.5 comes from this gene; cDNA EST yk461c1.5 comes from this gene [Caenorhabditis elegans] coil domains; cDNA EST yk302g12.5 comes from this gene; cDNA EST	0.096
241	U68138	Human PSD-95 mRNA, partial cds	2e-010	4521241	(AB024927) CsENDO-3 [Ciona savignyi]	2e-022
242	U88827	Aotus trivirgatus ribonuclease precursor gene, complete cds	6e-011	3914810	RIBONUCLEASE K6 PRECURSOR (RNASE K6) > gi 2745760 (AF037086) ribonuclease k6 precursor	1e-016
243	AF045573	Mus musculus FLI-LRR associated protein-1 mRNA, complete cds	2e-012	3025718	(AF045573) FLI-LRR associated protein-1 [Mus musculus]	3e-016

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
244	NM_001365.1	Homo sapiens discs, large (Drosophila) homolog 4 (DLG4) mRNA > :: gb U83192 HS U83192 Homo sapiens post-synaptic density protein 95 (PSD95) mRNA, complete cds	2e-012	4521241	(AB024927) CsENDO-3 [Ciona savignyi]	5e-020
245	U28049	Human TBX2 (TXB2) mRNA, complete cds.	7e-013	2501115	TBX2 PROTEIN (T-BOX PROTEIN 2)	2e-011
246	M23404	Chicken erythrocyte anion transport protein (band3) mRNA, complete cds.	2e-013	726403	(U23175) similar to anion exchange protein [Caenorhabditis elegans]	1e-025
247	AF005963	Homo sapiens XY homologous region, partial sequence	1e-014	104270	Ig heavy chain - clawed frog	1.9
248	M29863	Human farnesyl pyrophosphate synthetase mRNA	9e-015	182405	(M29863) farnesyl pyrophosphate synthetase [Homo sapiens]	0.005
249	D28126	Human gene for ATP synthase alpha subunit, complete cds (exon 1 to 12)	3e-015	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
250	Z80150	H.sapiens CACNL1A4 gene, exons 41 and 42 > :: emb A70716.1 A70716 Sequence 37 from Patent WO9813490	3e-015	3387914	(AF070550) cote1 [Homo sapiens]	3.5
251	U28049	Human TBX2 (TXB2) mRNA, complete cds.	4e-016	2501116	TBX2 PROTEIN (T-BOX PROTEIN 2) tbx gene [Mus musculus]	6e-009
252	U31629	Mus musculus C2C12 unknown mRNA, partial cds.	1e-017	3024998	HYPOTHETICAL HEART PROTEIN	3e-017
253	J05262	Human farnesyl pyrophosphate synthetase mRNA, complete cds.	1e-018	182405	(M29863) farnesyl pyrophosphate synthetase [Homo sapiens]	0.0001
254	D28126	Human gene for ATP synthase alpha subunit, complete cds (exon 1 to 12)	5e-019	<NONE>	<NONE>	<NONE>
255	D28126	Human gene for ATP synthase alpha subunit, complete cds (exon 1 to 12)	5e-019	3219984	HYPOTHETICAL PROTEIN MJ1597.1 region MJ1597.1 [Methanococcus jannaschii]	5.7

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
256	NM_004587.1	Homo sapiens ribosome binding protein 1 (dog 180kD homolog) (RRBP1) mRNA > :: gb AF006751 AF006751 Homo sapiens ES/130 mRNA, complete cds	2e-019	4759056	ribosome binding protein 1 (dog 180kD homolog) > gi 3299885 (AF006751) ES/130 [Homo sapiens]	0.004
257	U89915	Mus musculus junctional adhesion molecule (Jam) mRNA, complete cds	5e-020	3462455	(U89915) junctional adhesion molecule [Mus musculus]	2e-005
258	AF045573	Mus musculus FLI-LRR associated protein-1 mRNA, complete cds	5e-020	3025718	(AF045573) FLI-LRR associated protein-1 [Mus musculus]	9e-025
259	NM_004587.1	Homo sapiens ribosome binding protein 1 (dog 180kD homolog) (RRBP1) mRNA > :: gb AF006751 AF006751 Homo sapiens ES/130 mRNA, complete cds	2e-020	4759056	ribosome binding protein 1 (dog 180kD homolog) > gi 3299885 (AF006751) ES/130 [Homo sapiens]	0.0008

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
260	AF051098	Mus musculus seven transmembrane domain orphan receptor mRNA, complete cds	2e-021	3858883	(U67056) myosin I heavy chain kinase [Acanthamoeba castellanii] > gi 4206769 (AF104910) myosin I heavy chain kinase [Acanthamoeba castellanii]	0.002
261	AF051098	Mus musculus seven transmembrane domain orphan receptor mRNA, complete cds	2e-021	3858883	(U67056) myosin I heavy chain kinase [Acanthamoeba castellanii] > gi 4206769 (AF104910) myosin I heavy chain kinase [Acanthamoeba castellanii]	0.001
262	M13519	Human N-acetyl-beta-glucosaminidase (HEXB) mRNA, 3' end.	2e-021	4504373	hexosaminidase B (beta polypeptide) > gi 123081 sp P07686 HEXB_HUMAN BETA-HEXOSAMINIDASE BETA CHAIN PRECURSOR beta-N-acetylhexosaminidase (EC 3.2.1.52) beta chain - human > gi 386770 (M23294) beta-hexosaminidase beta-subunit [Homo sapiens]	2e-007

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
263	Z81014	Human DNA sequence from cosmid U65A4, between markers DXS366 and DXS87 on chromosome X *	2e-022	<NONE>	<NONE>	<NONE>
264	AF147311.1	Homo sapiens full length insert cDNA clone YA82F10	2e-022	3875904	(Z70207) predicted using Genefinder; similar to collagen; cDNA EST EMBL:D65905 comes from this gene; cDNA EST EMBL:D65858 comes from this gene; cDNA EST EMBL:D69306 comes from this gene; cDNA EST EMBL:D65755 comes from this gen...	0.07
265	AF037088	Gorilla gorilla ribonuclease k6 precursor, gene, complete cds	9e-024	3914791	RIBONUCLEASE K6 PRECURSOR (RNASE K6) > gi 2745752 (AF037082) ribonuclease k6 precursor	3e-019
266	Z81014	Human DNA sequence from cosmid U65A4, between markers DXS366 and DXS87 on chromosome X	8e-024	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
		*				
267	AF037088	Gorilla gorilla ribonuclease k6 precursor, gene, complete cds	9e-025	3914810	RIBONUCLEASE K6 PRECURSOR (RNASE K6) > gi 2745760 (AF037086) ribonuclease k6 precursor	4e-018
268	AF147311.1	Homo sapiens full length insert cDNA clone YA82F10	1e-026	131413	PULMONARY SURFACTANT-ASSOCIATED PROTEIN A PRECURSOR (SP-A) (PSP-A) (PSAP) precursor - rabbit > gi 165706 (J03542) apoprotein of surfactant [Oryctolagus cuniculus]	0.059
269	Z46786	D.melanogaster mRNA for acetyl-CoA synthetase	1e-027	1079042	acetyl-CoA synthetase - fruit fly	4e-025
270	NM_004039.1	Homo sapiens annexin II (lipocortin II) for lipocortin II, complete cds	4e-028	450448	(M33322) calpactin I heavy chain [Mus musculus]	0.1
271	X53064	Homo sapiens SPRR2A gene encoding small proline rich protein	1e-028	134846	SMALL PROLINE-RICH PROTEIN II rich protein [Homo sapiens]	0.005

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
272	M29863	Human farnesyl pyrophosphate synthetase mRNA	1e-028	4503685	farnesyl diphosphate synthase dimethylallyltranstransferase, geranyltranstransferase) bp313 to bp1374 is almost identical to human farnesyl pyrophosphate synthetase mRNA. [Homo sapiens]	2e-008
273	Z18950	H.sapiens genes for S100E calcium binding protein, CAPL, and S100D calcium binding protein EF-Hand patent US 5789248	5e-029	2493898	DOPAMINE-BETA-MONOOXYGENASE PRECURSOR (DOPAMINE BETA-HYDROXYLASE) (DBH) 1.14.17.1) precursor - mouse > gi 260873 bb s 119249 621 aa] [Mus sp.]	1.4
274	M19481	Human follistatin gene, exon 6.	5e-030	<NONE>	<NONE>	<NONE>

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
275	AF007155	Homo sapiens clone 23763 unknown mRNA, partial cds	2e-032	4502641	chemokine (C-C) receptor 7 TYPE 7 PRECURSOR (C-C CKR-7) (CC-CKR-7) (CCR-7) (MIP-3 BETA RECEPTOR) (EBV-INDUCED G PROTEIN-COUPLED RECEPTOR 1) (EBI1) (BLR2) > gi 1082381 p ir B55735 lymphocyte-specific G-protein-coupled receptor EBI1 - human > gi 468316 (L3158	1.6
276	M99624	Human epidermal growth factor receptor-related gene, 5' end.	8e-034	294845	(L13655) membrane protein [Saccharum hybrid cultivar H65-7052]	9e-014
277	U49082	Human transporter protein (g17) mRNA, complete cds	8e-035	1840045	(U49082) transporter protein [Homo sapiens]	1e-014

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
278	D50369	Homo sapiens mRNA for low molecular mass ubiquinone-binding protein, complete cds	9e-036	3024781	UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX UBIQUINONE-BINDING PROTEIN QP-C PROTEIN) (COMPLEX III SUBUNIT VII) ubiquinone-binding protein [Homo sapiens]	0.0002
279	AF086313	Homo sapiens full length insert cDNA clone ZD52B10	9e-036	2832777	(AL021086) /prediction=(method::; comes from the 5' UTR [Drosophila melanogaster]	1e-039
280	NM_004074.1	Homo sapiens cytochrome c oxidase subunit VIII (COX8), nuclear gene encoding mitochondrial protein, mRNA > :: gb J04823 HU MCOX8A Human cytochrome c oxidase subunit VIII (COX8) mRNA, complete cds.	1e-038	2499854	PROBABLE PEPTIDASE Y4SO > gi 2182630	2
281	AB024436.1	Homo sapiens mRNA for beta-1,4-galactosyltransferase IV, complete cds	2e-041	3132900	(AF038662) beta-1,4-galactosyltransferase [Homo sapiens] beta-1,4-galactosyltransferase	4e-016

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
					galactosyltransferase IV [Homo sapiens]	
282	AF057734	Homo sapiens 17-beta-hydroxysteroid dehydrogenase IV (HSD17B4) gene, exon 16	2e-043	2842416	(AL008730) dJ487J7.1.1 (putative protein dJ487J7.1 isoform 1) [Homo sapiens]	3e-062
283	Z69650.1	Human DNA sequence from cosmid L69F7B, Huntington's Disease Region, chromosome 4p16.3 contains Huntington Disease (HD) gene	2e-044	1872200	(U22376) alternatively spliced product using exon 13A	1e-008
284	NM_003938.1	Homo sapiens adaptin, delta (ADTD) mRNA > :: gb U91930 HS U91930 Homo sapiens AP-3 complex delta subunit mRNA, complete cds	2e-044	3478639	(AC005545) delta-adaptin, partial CDS [Homo sapiens]	3e-016
285	AF026029	Homo sapiens poly(A) binding protein II (PABP2) gene, complete cds	8e-045	1916930	(U88570) CREB-binding protein homolog [Drosophila melanogaster]	7.6
286	AB006622	Homo sapiens mRNA for KIAA0284 gene, partial cds	1e-045	73404	E2 protein - human papillomavirus type 5	0.11

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
287	U90918	Human clone 23654 mRNA sequence	1e-048	3877568	(Z70208) similar to collagen	0.042
288	AB006622	Homo sapiens mRNA for KIAA0284 gene, partial cds	1e-049	73404	E2 protein - human papillomavirus type 5	0.11
289	AL049258.1	Homo sapiens mRNA; cDNA DKFZp564E173 (from clone DKFZp564E173)	1e-050	<NONE>	<NONE>	<NONE>
290	AF022367	Homo sapiens beta-1,4-galactosyltransferase mRNA, complete cds	5e-051	3132900	(AF038662) beta-1,4-galactosyltransferase [Homo sapiens] beta-1,4-galactosyltransferase IV [Homo sapiens]	6e-019
291	AF057734	Homo sapiens 17-beta-hydroxysteroid dehydrogenase IV (HSD17B4) gene, exon 16	7e-053	2842416	(AL008730) dJ487J7.1.1 (putative protein dJ487J7.1 isoform 1) [Homo sapiens]	6e-055
292	AF097709	Homo sapiens serine protease (PRSS11) mRNA, partial cds	8e-055	4506141	protease, serine, 11 (IGF binding) > gi 1513059 d bj BAA13322 (D87258) serin protease with IGF-binding motif [Homo sapiens] protease, PRSS11 [Homo sapiens]	2e-017

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
293	U31629	Mus musculus C2C12 unknown mRNA, partial cds.	9e-057	3025215	HYPOTHETICAL 81.0 KD PROTEIN C35D10.4 IN CHROMOSOME III > gi 2146877 pir S72572 probable ABC1 protein homolog - Caenorhabditis elegans protein (Swiss-Prot Acc: P27697) [Caenorhabditis elegans]	5e-033
294	AB006622	Homo sapiens mRNA for KIAA0284 gene, partial cds	8e-057	73404	E2 protein - human papillomavirus type 5	1.7
295	AF025439	Homo sapiens Opa-interacting protein OIP3 mRNA, partial cds	4e-059	<NONE>	<NONE>	<NONE>
296	M99624	Human epidermal growth factor receptor-related gene, 5' end.	1e-060	123364	SEGMENTATION PROTEIN EVEN-SKIPPED fly (Drosophila sp.) > gi 157387 (M14767) even-skipped gene [Drosophila melanogaster]	5.3
297	AF045573	Mus musculus FLI-LRR associated protein-1 mRNA, complete cds	5e-061	3025718	(AF045573) FLI-LRR associated protein-1 [Mus musculus]	7e-029
298	AB006622	Homo sapiens mRNA for KIAA0284 gene, partial cds	2e-062	2119133	ribosomal protein S17 - cat (fragment) musculus]	2e-015

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
299	M30702	Human amphiregulin (AR) gene, exon 5, clones lambda-ARH(6,12).	2e-063	4502199	amphiregulin (schwannoma-derived growth factor) > gi 113754 sp P15514 AMP_R_HUMAN AMPHIREGULIN PRECURSOR (AR) (COLORECTUM CELL-DERIVED GROWTH FACTOR) (CRDGF) > gi 107391 pir A34702 amphiregulin precursor - human > gi 178890 (M30703) amphiregulin [Homo sapien	0.0002
300	L38847	Mus musculus hepatoma transmembrane kinase ligand Sequence 1 from patent US 5624899	6e-064	3861228	(AJ235272) unknown [Rickettsia prowazekii]	2.9
301	L38847	Mus musculus hepatoma transmembrane kinase ligand Sequence 1 from patent US 5624899	6e-064	3861228	(AJ235272) unknown [Rickettsia prowazekii]	2.9
302	Z78141	M.musculus partial cochlear mRNA (clone 29C9)	8e-066	1490324	(Z78141) unknown [Mus musculus]	8e-019

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
303	X12650	Mus musculus gene for beta-tropomyosin	2e-072	833602	(X54277) cardiac tropomyosin [Coturnix coturnix]	7e-022
304	M87635	Mouse beta-tropomyosin 2 mRNA, complete cds.	2e-084	1216293	(L35239) cardiac tropomyosin [Xenopus laevis]	5e-019
305	M13364	Rabbit calcium-dependent protease, small subunit mRNA, complete cds.	2e-084	115611	CALCIUM-DEPENDENT PROTEASE, SMALL NEUTRAL PROTEINASE) (CANP) > gi 108563 pir A34466 calpain (EC 3.4.22.17) II light chain - bovine 3.4.22.17) [Bos taurus]	1e-058
306	M87635	Mouse beta-tropomyosin 2 mRNA, complete cds.	3e-088	1216293	(L35239) cardiac tropomyosin [Xenopus laevis]	9e-028
307	M87635	Mouse beta-tropomyosin 2 mRNA, complete cds.	5e-092	1216293	(L35239) cardiac tropomyosin [Xenopus laevis]	2e-035
308	X85992	M.musculus mRNA for semaphorin C	8e-097	2137756	semaphorin C - mouse (fragment) musculus]	2e-048

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
309	M24103	Bovine ADP/ATP translocase T2 mRNA, complete cds.	e-103	113463	ADP,ATP CARRIER PROTEIN, LIVER ISOFORM T2 (ADP/ATP TRANSLOCASE 3) (ADENINE NUCLEOTIDE TRANSLOCATOR 3) (ANT 3) > gi 86757 pir S03894 ADP,ATP carrier protein T2 - human	2e-035
310	U48852	Cricetulus griseus HT protein mRNA, complete cds.	e-107	1216486	(U48852) HT protein [Cricetulus griseus]	3e-057
311	X76168	R.norvegicus mRNA for connexin 30.3	e-112	544118	GAP JUNCTION BETA-5 PROTEIN (CONNEXIN 30.3) (CX30.3) > gi 481577 pir S38891 connexin 30.3 - rat > gi 431204 emb CAA53762 (X76168) connexin 30.3	1e-063
312	X76168	R.norvegicus mRNA for connexin 30.3	e-115	461864	GAP JUNCTION BETA-5 PROTEIN junction protein Cx30.3 - mouse > gi 192647 (M91443) connexin 30.3 [Mus musculus]	7e-064

Table 2

	Nearest Neighbor (BlastN vs. Genbank)			Nearest Neighbor (BlastX vs. Non-Redundant Proteins)		
313	AJ009634.1	Mus musculus fjx1 gene	e-137	4138203	(AJ009634) Fjx1 [Mus musculus]	5e-065
314	X76168	R.norvegicus mRNA for connexin 30.3	e-130	544118	GAP JUNCTION BETA-5 PROTEIN (CONNEXIN 30.3) (CX30.3) >gi 481577 pi r S38891 connexin 30.3 - rat >gi 431204 e mb CAA53762 (X76168) connexin 30.3	2e-074

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
4	819498					
8	728115	21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
		15,16 (Normal Colon vs. Colon Tumor)	0	7		6.62
		16,17 (Colon Tumor vs. Colon Metastasis)	7	0	7.11	
9	372700					
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	3	50		11.93
		19,20 (Colon Tumor vs. Colon Tumor Metastasis)	8	0	5.98	
12	729832					
		15,16 (Normal Colon vs. Colon Tumor)	0	11		10.41
		16,17 (Colon Tumor vs. Colon Metastasis)	11	0	11.17	
13	505514					
		23,24 (Normal Lung vs. Lung Tumor)	26	10	2.63	
17	549934					
		21,22 (Normal Prostate vs. Cancerous Prostate)	8	0	7.87	
		16,17 (Colon Tumor vs. Colon Metastasis)	3	20		6.56
		15,16 (Normal Colon vs. Colon Tumor)	11	3	3.88	
25	450399					

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
		15,16 (Normal Colon vs. Colon Tumor)	28	68		2.3
		15,17 (Normal Colon vs. Colon Metastasis)	28	117		3.89
26	450982					
		16,17 (Colon Tumor vs. Colon Metastasis)	14	32		2.25
28	379302					
		21,22 (Normal Prostate vs. Cancerous Prostate)	8	1	7.87	
43	817503					
		21,22 (Normal Prostate vs. Cancerous Prostate)	18	4	4.43	
48	830085					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	9		9.15
52	830931					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	7		7.12
55	819046					
		21,22 (Normal Prostate vs. Cancerous Prostate)	2	13		6.61
58	728115					
		15,16 (Normal Colon vs. Colon Tumor)	0	7		6.62
		16,17 (Colon Tumor vs. Colon Metastasis)	7	0	7.11	
65	553242					

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
71	820061	_16,17 (Colon Tumor vs. Colon Metastasis)	0	6		5.91
		_21,22 (Normal Prostate vs. Cancerous Prostate)	1	20		20.33
78	220584					
		_08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	1	12		8.59
80	549934					
		_16,17 (Colon Tumor vs. Colon Metastasis)	3	20		6.56
		_15,16 (Normal Colon vs. Colon Tumor)	11	3	3.88	
		_21,22 (Normal Prostate vs. Cancerous Prostate)	8	0	7.87	
86	819460					
		_21,22 (Normal Prostate vs. Cancerous Prostate)	18	1	17.7	
95	551785					
		_21,22 (Normal Prostate vs. Cancerous Prostate)	0	6		6.1
96	17092					
		_03,04 (Breast, High Metastatic Potential vs. Breast, Non-Metastatic)	0	25		25.62
99	745559					
		_21,22 (Normal Prostate vs. Cancerous Prostate)	1	9		9.15

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
101	379879					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	9		9.15
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	0	13		9.3
107	268290					
		21,22 (Normal Prostate vs. Cancerous Prostate)	33	69		2.13
108	818043					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
114	450247					
		21,22 (Normal Prostate vs. Cancerous Prostate)	23	8	2.83	
115	819273					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
116	587779					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
118	615617					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	7		7.12
121	818682					
		21,22 (Normal Prostate vs. Cancerous Prostate)	11	2	5.41	

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
123	484413					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
124	819273					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
127	818682					
		21,22 (Normal Prostate vs. Cancerous Prostate)	11	2	5.41	
131	819273					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
147	820061					
		21,22 (Normal Prostate vs. Cancerous Prostate)	1	20		20.33
153	375958					
		21,22 (Normal Prostate vs. Cancerous Prostate)	2	11		5.59
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	0	9		6.44
155	831049					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	11		11.18
157	553200					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	6		6.1

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
158	139677					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
159	139677					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
163	375958					
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	0	9		6.44
		21,22 (Normal Prostate vs. Cancerous Prostate)	2	11		5.59
168	831812					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	7		7.12
176	193373					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
177	400619					
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	6	0	8.38	
178	831149					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	7		7.12
180	817503					
		21,22 (Normal Prostate vs. Cancerous Prostate)	18	4	4.43	

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
187	648679					
		23,24 (Normal Lung vs. Lung Tumor)	11	1	11.11	
		16,17 (Colon Tumor vs. Colon Metastasis)	79	0	80.23	
		15,17 (Normal Colon vs. Colon Metastasis)	7	0	7.51	
		15,16 (Normal Colon vs. Colon Tumor)	7	79		10.68
190	373928					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
195	373928					
		21,22 (Normal Prostate vs. Cancerous Prostate)	7	0	6.88	
198	372700					
		19,20 (Colon Tumor vs. Colon Tumor Metastasis)	8	0	5.98	
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	3	50		11.93
204	379105					
		15,16 (Normal Colon vs. Colon Tumor)	0	8		7.57
205	831188					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	8		8.13
209	831812					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	7		7.12

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
213	831026					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	10		10.17
215	380207					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	6		6.1
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	0	8		5.72
216	819460					
		21,22 (Normal Prostate vs. Cancerous Prostate)	18	1	17.7	
224	819201					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
225	374826					
		15,17 (Normal Colon vs. Colon Metastasis)	5	20		3.73
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	38	132		2.49
		15,16 (Normal Colon vs. Colon Tumor)	5	18		3.41
231	553242					
		16,17 (Colon Tumor vs. Colon Metastasis)	0	6		5.91
246	220584					
		08,09 (Lung, High Metastatic Potential vs. Lung, Low Metastatic Potential)	1	12		8.59
248	819498					

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
253	819498					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	
256	831160					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	12		12.2
259	831160					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	12		12.2
262	373298					
		15,17 (Normal Colon vs. Colon Metastasis)	126	42	3.22	
		15,16 (Normal Colon vs. Colon Tumor)	126	59	2.26	
270	450262					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	8		8.13
271	484703					
		21,22 (Normal Prostate vs. Cancerous Prostate)	28	0	27.54	
272	819498					
		21,22 (Normal Prostate vs. Cancerous Prostate)	6	0	5.9	

Table 4

SEQ ID	CLUST	PairAB-text	CLONES in A	CLONES in B	RATIO PLUS	RATIO MINUS
273	406043					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	6		6.1
274	817500					
		21,22 (Normal Prostate vs. Cancerous Prostate)	2	18		9.15
275	818180					
		21,22 (Normal Prostate vs. Cancerous Prostate)	2	10		5.08
280	429009					
		21,22 (Normal Prostate vs. Cancerous Prostate)	8	1	7.87	
284	383021					
		21,22 (Normal Prostate vs. Cancerous Prostate)	3	12		4.07
289	831580					
		21,22 (Normal Prostate vs. Cancerous Prostate)	0	6		6.1
311	763446					
		21,22 (Normal Prostate vs. Cancerous Prostate)	11	1	10.82	
312	763446					
		21,22 (Normal Prostate vs. Cancerous Prostate)	11	1	10.82	
314	763446					
		21,22 (Normal Prostate vs. Cancerous Prostate)	11	1	10.82	
315	10154					
		3,4 (Breast,High Metastatic Potential vs. Breast, Low Metastatic)	3	317		108.1

Table 7

Library No.	Clones		
es75	M00063947D:D01	es79	M00064003B:C10
	M00063158A:A01		M00064302A:D10
	M00063517A:A04		M00064309C:H09
	M00063520D:E11		M00064310D:F03
	M00063638C:G12		M00064322C:A10
	M00063642B:A08		M00064359B:H12
	M00063686B:E07		M00064390A:C05
	M00063689D:E12		M00064404A:B05
	M00063781B:B10		M00064404C:G05
	M00063826A:D03		M00064404D:A06
es76	M00063838B:G08	es80	M00064429D:B07
	M00063838B:G08		M00064446A:D11
	M00063841A:B09		M00064457D:C09
	M00063886A:B06		M00064476D:C04
	M00063910D:A12		M00064506A:C07
	M00063912A:D06		M00064514A:G10
	M00063920D:H05		M00064520A:F08
	M00063928A:G09		M00064579D:E11
	M00063934B:E04		M00064620C:D01
	M00063945A:C03		M00064624D:C09
es77	M00064032D:G04	es81	M00064633C:A03
	M00064046A:G02		M00064637B:F03
	M00064053C:G04		M00064690A:C04
	M00064053D:F02		M00064690A:C04
	M00064082A:A08		M00064714A:G03
	M00064089B:F09		M00064723D:H11
	M00064132B:B07		GKC10154-1
	M00064138A:F11		GKC10154-3
	M00064161B:G04		
	M00064175B:B09		
es78	M00064178C:C04		
	M00064179A:C04		
	M00064200D:E08		
	M00064248A:E02		
	M00064270B:B03		
	M00064271B:D03		
	M00063580C:A06		
	M00063594B:H07		
	M00064002C:F06		
	M00064002C:H09		

es82	M00063151A:G06	M00063852D:F07
	M00063151D:B10	M00063888D:D05
	M00063152C:B07	M00063888D:F02
	M00063156D:H10	M00063890A:F11
	M00063158A:E11	M00063890A:H04
	M00063158A:E11	M00063891A:F11
	M00063452A:F08	M00063892B:G02
	M00063453B:F08	M00063898A:A10
	M00063462D:D07	M00063915C:E01
	M00063463D:B05	M00063919C:E07
	M00063466C:C11	M00063920D:H02
	M00063467D:H07	M00063922B:A12
	M00063478C:D01	M00063925B:F04
	M00063482A:A08	M00063926A:H04
	M00063482A:F07	M00063931B:E10
	M00063485A:E05	M00063931B:F07
	M00063487C:C02	M00063932D:G08
	M00063514C:D03	M00063934C:C10
	M00063514C:E08	M00063938B:H07
	M00063515B:F06	M00063939C:D06
	M00063515B:H02	M00063939C:H01
	M00063518D:A01	M00063940D:F09
	M00063520D:D08	M00063940D:F09
	M00063604A:B11	M00063941B:C12
	M00063606C:B04	M00063943B:G12
	M00063610D:C11	M00063949D:A05
	M00063613D:C11	M00064021D:H01
	M00063617D:F09	M00064025D:E07
	M00063627C:F06	M00064025D:H12
	M00063636A:E01	M00064033C:C11
	M00063681B:C02	M00064033D:B01
	M00063682A:C04	M00063843B:D07
	M00063685A:C02	M00063848C:G11
	M00063774A:D09	M00063852B:D08
	M00063784A:H12	M00063818C:A09
	M00063784C:E10	M00063828A:H12
	M00063785C:F03	M00063828D:E05
	M00063795C:D09	M00063839A:F01
	M00063801B:D04	M00063841A:E08
	M00063804C:A11	
	M00063805D:E05	
	M00063807A:D12	
	M00063810C:E03	

es83

M00064043D:C09	M00063577C:C02
M00064048C:G12	M00063578B:E02
M00064053B:D09	M00063578C:A06
M00064057C:H10	M00063580D:B06
M00064059A:C11	M00063593A:D03
M00064060B:D03	M00063600C:C09
M00064079C:A10	M00063955C:F07
M00064082D:D10	M00063955D:F05
M00064083D:E05	M00063956A:F05
M00064086C:E01	M00063957A:E02
M00064090C:A02	M00063957A:E02
M00064090D:D09	M00063967C:A12
M00064105B:A03	M00063967D:G02
M00064106C:G03	M00063968D:G08
M00064113B:C04	M00063972C:E10
M00064115B:E12	M00063978B:B06
M00064119B:H10	M00063981D:A06
M00064119C:D12	M00063990A:D05
M00064122C:B06	M00063990A:D05
M00064126C:C02	M00063997C:B12
M00064126C:F12	M00063998C:E09
M00064136C:D12	M00064000B:C03
M00064144D:A07	M00064001A:B03
M00064151B:C07	M00064005D:A08
M00064159A:H03	M00064008A:B01
M00064165A:B12	M00064009A:C01
M00064171D:E05	M00064014D:H05
M00064171D:E05	M00064018C:E07
M00064172C:A02	M00064293D:B12
M00064173B:E01	M00064294D:F01
M00064176D:H10	M00063557D:C07
M00064178B:A05	M00063559D:G03
M00064178B:A05	M00063571B:G03
M00064180A:G03	M00063575B:G02
M00064186C:B03	M00063555B:D01
M00064188B:G08	M00063533A:C12
M00064194C:D02	M00063534C:A02
M00064212D:E04	M00063538D:B01
M00064260C:E05	M00063539C:C11
M00064268D:G03	
M00064272C:G01	
M00063163A:G04	
M00063165A:C09	

es84	M00064307B:G02	M00064564A:C02
	M00064307C:G03	M00064568A:H06
	M00064310C:A10	M00064569B:A09
	M00064328B:H04	M00064569B:A09
	M00064328B:H09	M00064571C:C04
	M00064337D:F01	M0064577C:B120
	M00064341A:C02	M00064579A:C06
	M00064345A:A03	M00064593A:A05
	M00064346C:B09	M00064593D:C01
	M00064349D:H01	M00064601C:G07
	M00064352C:H01	M00064601D:B05
	M00064354A:A10	M00064605C:G05
	M00064358A:G03	M00064610D:H01
	M00064358C:D09	M00064620D:G05
	M00064375B:G07	M00064624C:B03
	M00064376A:A05	M00064631A:C07
	M00064385D:C11	M00064631A:C07
	M00064386B:C02	M00064631C:H11
	M00064386B:C02	M00064636B:A04
	M00064393B:H04	M00064649A:E04
	M00064399A:E01	M00064650B:B07
	M00064405B:C04	M00064652B:D09
	M00064406B:H06	M00064675C:E09
	M00064414D:D06	M00064678D:F05
	M00064415B:G03	M00064693D:F08
	M00064424B:C12	M00064723C:H04
	M00064428B:A12	M00064723D:H03
	M00064447B:A07	M00064723D:H03
	M00064447B:C06	M00003773D:H02
	M00064450C:E07	M00021929A:D03
	M00064452D:E11	M00043134A:A05
	M00064454A:H10	M00064534D:F06
	M00064454C:B06	M00064550A:A07
	M00064460C:B01	M00064554D:A03
	M00064467B:D06	M00064526D:F05
	M00064481C:F03	M00064527A:H07
	M00064508A:B09	M00064530B:H02
	M00064514D:F11	M00064532D:G06
	M00064517B:F04	M00064520A:E04
	M00064517B:F10	M00064520A:E04
	M00064517C:F11	M00064524A:A09

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
15	21	Ascending colon ₁	4.0	T3	G2	extending into subserosal adipose tissue	positive	3/8	N1	negative		MX	invasive adenocarcin oma, moderately differentiate d; focal perineural invasion is seen
52	71	Ascending colon	9.0	T3	G3	Invasion through muscularis propria, subserosal involvement; ileocec. valve involvement	negative	0/12	N0	negative		M0	Hyperplastic polyp in appendix.
121	140	Sigmoid	6	T4	G2	Invasion of muscularis propria into serosa, involving submucosa of urinary bladder	negative	0/34	N0	negative		M0	Perineural invasion; donut anastomosis negative. One tubulovillou s and one tubular adenoma with no high

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
													grade dysplasia.
125	144	Cecum	6	T3	G2	Invasion through the muscularis propria into subserosal adipose tissue. Ileocecal junction.	negative	0/19	N0	negative		M0	patient history of metastatic melanoma
128	147	Transverse colon	5.0	T3	G2	Invasion of muscularis propria into pericolonic fat	positive	1/5	N1	negative		M0	
130	149	Splenic flexure	5.5	T3		through wall and into surrounding adipose tissue	positive	10/24	N2	negative		M1	

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
133	152	Rectum	5.0	T3	G2	Invasion through muscularis propria into non- peritonealized pericolonic tissue; gross configuration is annular.	negative	0/9	N0	negative		M0	Small separate tubular adenoma (0.4 cm)
141	160	Cecum	5.5	T3	G2	Invasion of muscularis propria into pericolonic adipose tissue, but not through serosa. Arising from tubular adenoma.	positive	7/21	N2	positive (Liver)	adenoca rcinoma consista nt with primary	M1	Perineural invasion identified adjacent to metastatic adenocarcin oma.

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
156	175	Hepatic flexure	3.8	T3	G2	Invasion through mucularis propria into subserosa/peric olic adipose, no serosal involvement. Gross configuration annular.	positive	2/13	N1	negative		M0	Separate tubulovillou s and tubular adenomas
228	247	Rectum	5.8	T3	G2 to G3	Invasion through mucularis propria to involve subserosal, perirectoal adipose, and serosa	positive	1/8	N1	negative		MX	Hyperplastic polyps

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
264	283	Ascending colon	5.5	T3	G2	Invasion through muscularis propria into subserosal adipose tissue.	negative	0/10	N0	negative		M0	Tubulovillous adenoma with high grade dysplasia
266	285	Transverse colon	9	T3	G2	Invades through muscularis propria to involve pericolonic adipose, extends to serosa.	negative	0/15	N1	positive (Mesenteric deposit)	0.4 cm, may represent lymph node completely replaced by tumor	MX	
268	287	Cecum	6.5	T2	G2	Invades full thickness of muscularis propria, but mesenteric adipose free of malignancy	negative	0/12	N0	negative		M0	

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
278	297	Rectum	4	T3	G2	Invasion into perirectal adipose tissue.	positive	7/10	N2	negative		M0	Descending colon polyps, no HGD or carcinoma identified..
295	314	Ascending colon	5.0	T3	G2	Invasion through muscularis propria into percolic adipose tissue.	negative	0/12	N0	negative		M0	Melanosis coli and diverticular disease.
339	358	Rectosigmoid	6	T3	G2	Extends into perirectal fat but does not reach serosa	negative	0/6	N0	negative		M0	1 hyperplastic polyp identified
341	360	Ascending colon	2 cm invasive	T3	G2	Invasion through muscularis propria to involve pericolonic fat. Arising from villous adenoma.	negative	0/4	N0	negative		MX	

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
356	375	Sigmoid	6.5	T3	G2	Through colon wall into subserosal adipose tissue. No serosal spread seen.	negative	0/4	N0	negative		M0	
360	412	Ascending colon	4.3	T3	G2	Invasion thru muscularis propria to pericolonic fat	positive	1/5	N1	negative		M0	Two mucosal polyps
392	444	Ascending colon	2	T3	G2	Invasion through muscularis propria into subserosal adipose tissue, not serosa.	positive	1/6	N1	positive (Liver)	Macro- vesicular and micro- vesicular steatosis	M1	Tumor arising at prior ileocolic surgical anastomosis.

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc Met	Descrip Distant Met	Dist Met Grade	Comment
393	445	Cecum	6.0	T3	G2	Cecum, invades through muscularis propria to involve subserosal adipose tissue but not serosa.	negative	0/21	N0	negative		M0	
413	465	Ascending colon	4.8	T3	G2	Invasive through muscularis to involve periserosal fat; abutting ileocecal junction.	negative	0/7	N0	positive (Liver)	adenoca rcinoma in multiple slides	M1	redagnosis of oophorecto my path to metastatic colon cancer.

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
						Invasion through muscularis propria involving pericolic adipose, serosal surface uninvolved					moderately differentiated adenocarcinoma, consistent with primary		Anatomical location of primary not notated in report. Evidence of chronic colitis.
505	383		7.5 cm max dim	T3	G2		positive	2/17	N1	positive (Liver)		M1	
517	395	Sigmoid	3	T3	G2	penetrates muscularis propria, involves pericolic fat.	positive	6/6	N2	negative		M0	No mention of distant met in report

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode e Met	Regional Lymphnode e Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
534	553	Ascending colon	12	T3	G3	Invasion through the muscularis propria involving pericolic fat. Serosa free of tumor.	negative	0/8	N0	negative		M0	Omentum with fibrosis and fat necrosis. Small bowel with acute and chronic serositis, focal abscess and adhesions.
546	565	Ascending colon	5.5	T3	G2	Invasion through muscularis propria extensively through submucosal and extending to serosa.	positive	6/12	N2	positive (Liver)	metastatic adenocarcinoma	M1	

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
577	596	Cecum	11.5	T3	G2	Invasion through the bowel wall, into suberosal adipose. Serosal surface free of tumor.	negative	0/58	N0	negative		M0	Appendix dilated and fibrotic, but not involved by tumor
695	714	Cecum	14	T3	G2	extending through bowel wall into serosal fat	negative	0/22	N0	negative		MX	tubular adenoma and hyperplastic polyps present, moderately differentiated adenoma with mucinous differentiation (% not stated)
784	803	Ascending colon	3.5	T3	G3	through muscularis propria into pericolonic soft tissues	positive	5/17	N2	positive (Liver)		M1	invasive poorly differentiated adenosquamous

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode Met	Regional Lymphnode Grade	Distant Met & Loc	Descrip Distant Met	Dist Met Grade	Comment
													carcinoma
786	805	Descending colon	9.5	T3	G2	through muscularis propria into pericolic fat, but not at serosal surface	negative	0/12	N0	positive (Liver)		M1	moderately differentiated invasive adenocarcinoma
791	810	Ascending colon	5.8	T3	G3	through the muscularis propria into pericolic fat	positive	13/25	N2	positive (Liver)		M1	poorly differentiated invasive colonic adenocarcinoma

Table 8 Patient ID	Path Report ID	Anatomical Loc	Primary Tumor Size	Primary Tumor Grade	Histopath Grade	Local Invasion	Lymphnode Met	Incidence Lymphnode e Met	Regional Lymphnode e Grade	Distant Met & Loc Met	Descrip Distant Met	Dist Met Grade	Comment
888	908	Ascending colon	2.0	T2	G1	into muscularis propria	positive	3/21	N0	positive (Liver)		M1	well- to moderately- differentiate d adenocarcin oma; this patient has tumors of the ascending colon and the sigmoid colon
889	909	Cecum	4.8	T3	G2	through muscularis propria int subserosal tissue	positive	1/4	N1	positive (Liver)		M1	moderately differentiate d adenocarcin oma

We Claim:

1. An isolated polynucleotide comprising a nucleotide sequence which hybridizes under stringent conditions to a sequence selected from the group consisting of SEQ ID NOS: 1-316.

2. An isolated polynucleotide comprising at least 15 contiguous nucleotides of a nucleotide sequence having at least 90% sequence identity to a sequence selected from the group consisting of: SEQ ID NOS:1-316, a degenerate variant of SEQ ID NOS:1-316, an antisense of SEQ ID NOS:1-316, and a complement of SEQ ID NOS:1-316.

3. A polynucleotide comprising a nucleotide sequence of an insert contained in a clone deposited as clone number xx of ATCC Deposit Number xx .

4. An isolated cDNA obtained by the process of amplification using a polynucleotide comprising at least 15 contiguous nucleotides of a nucleotide sequence of a sequence selected from the group consisting of SEQ ID NOS:1-316.

5. The isolated cDNA of claim 4, wherein amplification is by polymerase chain reaction (PCR) amplification.

6. An isolated recombinant host cell containing the polynucleotide according to claims 1, 2, 3, or 4.

7. An isolated vector comprising the polynucleotide according to claims 1, 2, 3, or 4.

8. A method for producing a polypeptide, the method comprising the steps of: culturing a recombinant host cell containing the polynucleotide according to claims 1, 2, 3, or 4, said culturing being under conditions suitable for the expression of an encoded polypeptide; recovering the polypeptide from the host cell culture.

9. An isolated polypeptide encoded by the polynucleotide according to claims 1, 2, 3, or 4.

10. An antibody that specifically binds the polypeptide of claim 9.

11. A method of detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, the method comprising the step of:

detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous, where the gene product is encoded by a gene comprising an identifying sequence of at least one of SEQ ID NOS:1-316;

5 wherein detection of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived.

12. A library of polynucleotides, wherein at least one of the polynucleotides comprises the sequence information of the polynucleotide according to claims 1, 2, 3, or 4

10 13. The library of claim 12, wherein the library is provided on a nucleic acid array.

14. The library of claim 12, wherein the library is provided in a computer-readable format.

15 15. A method of inhibiting tumor growth by modulating expression of a gene product, the gene product being encoded by a gene identified by a sequence selected from the group consisting of SEQ ID NOS:1-316.

1
1 GCCAGATTCCGGCAGAGGGGGCTAGGACCCCTCCGAGCCAGGTGTGGGATATAGTCTCGTGGTCCGCCGTTTTTTAAGCCG 80

1
81 GTCTGAAAAGCGCAATATTCCGGTGGGAGTGACCCGATTTTCAGGCTGCTATCCATGTCCAGGGCCAAACATGAATCCT 160

1
161 ATTGCTCTTGGGAGCCGCTGGCTTGCTTATGCAGAAAACAAGTTGATTGATGTCTATCAGTCCCGTGOTGGAGCCCTGTGG 240

1
241 AGACAAACATTCACTCTTATACTGCCACAGTCATTAGTGTGCTAAACATTTAAAAGTGCCCTGACAATGTAGGGAAAG 320

1
321 TGGTGAATCAGCTGACAGGCACACTGCCCTTCAGGTGTACAGAGATGATGTTGCCATCCACAGTAATTCACGGCGGAGT 400

1
401 CCTTTGGTCCCAGGCATCATCACAGTTATTGACACCGAAACCGTTGGAGAGGGCCAGGTGCTTGTGAGTGAGGATTCTGA 480

1
481 CAGTGATGGCATTTGTGGCCCACTTCCCTGCCCATGAGAAGCCAGTGCTGTGCATGGCTTTAATACAAAGTGAATGCTTC 560

1
561 TAGTCACAACAGACACCCCTGGCCATGACTTTCATGTCTTCCAAATTCGACTCATCCTTGGTCCCTCATCACAAATGTCT 640

1
641 GTCCACCATCTGTATACTCTTCACAGGGGAGAACTGAAGCCAAAGTACAGGACATCTGCTTCAGCCATGACTGTGGCTG 720

1
721 GGTGTGGTCACTACTCTCCGGGTACTTCCCAGGTTTTCCCACATCAACCTTATGGTGGCCAGCCTTGTGTGTGTACAC 800

1
801 ATATGTCAACCAAGTAGTAGTGAATCGCATGAGCCGTTTCCAGAAAAGTGTGGACTGGAGAGATTGAACAAAGACTGACG 880

1
881 TCTAAGCAAGGAGGTGGCTGTAGCCCTGTTCCAGGTCTATCAAGCAGCCCTTCTGGGTCAACCTTGCATGGGAAACTGAA 960

1
961 CAGCCAAAGACTCCTATAACAATTTACCAACAACCCCTGGCAACCCTCGGCTCTCTCTCTTCCCAGCTTGATGGTAG 1040

1
1041 TGATGCCCTCTTGCACAAATCAAGCAGCCAATGACATGGGGACCATCACCAACGAACCGGGCCTTATCTCTTTGGAGCG 1120

1
1121 GGGTGTTTTTCCATAAAGCCCCATGCAAGTTAAACCTCTCCACAAATTCACCCAGCAAATTGATGGGCGGAGAAATT 1200

1
1201 TTGTGTGGCTGCTATCTTCGAACATCCAGGTCAATGTTTGCAAATAATGCAGGTCTGAAAACAGAAAAGATCAGTCCA 1280

FIGURE 1A

1 0
1281 AACAAAGTTGTAGTTGAGTCCCTGTACATTATCAGTTGCTATGGCACCTTAGTGGAACACATGATGGAGCCCGACCCCTC 1360

1 0
1361 AGCACTGCACCCCAAGATTAGTGACGACACACCACTGGAAATGATGACATCGCCTCGAGCCAGCTGGACTCTGOTTAGAAC 1440

1 0
1441 CCCTCAATGGAATGAATTGCAGCCACCGTTTAAATGCAAAACACCCCTCTGCTCCTCGCTGCAGATGCAGTACAGTATTATC 1520

1 0
1521 AGTTCTGCTTGCTGGCCTGGTTCCCCCTGGAAGTCTGGGGCCCATTAAGTCCACATGGGTCTTACGACAGTTTAGCTTCT 1600

1 0
1601 GACCATAGTGGACAGGAAGATGAAGAATGGCTTCCAGGTTGAAATTGTAACACACACTGGACCCCATAGACGTCTGTG 1680

1 0
1681 GATGGGTCCACAGTTCCAGTTCAAAACCATCCATCCCTCAGGCCAAACCACAGTTATCTCATCCAGTTCATCTGTGTTC 1760

1 0
1761 AGTCTCATGGTCCGAGTGACAGCCACAGCCTCTTTGGATTTTGATACAGATGATCTTGATCTCAACAGTCTCAGGATC 1840

1 50
1841 CAGCCAGTCCGCTCTGACCCCGTCAGCATGCCAGGGTCATCCCTGCCAGTCTCTGATCGAAGGGGAGTTTCCACAGTGAT 1920
TCGAGCCCATGCCAGCGGGGGCGGTTAGCTCGGGCTCTTCTGACCCC
* * * * *

51 130
1911 CGATCCTGGGGCCGAGGTACCTTTGACAGGAGCGTGACCTGCTGGAGGTGTGCGGGAGCTGGCCCTGAGGGCTTCGGGCT 1998
TGATGCTG--CCTCAGGTACCTTTGACAGGAGCGTGACCTGCTGGAGGTGTGCGGGAGCTGGCCCTGAGGGCTTCGGGCT
* * * * *

131 210
1999 GCGGCACATGTCCTCCATGGAGCACACGGAGGAGGGGCTTCGGGGAGCGACTTGCCGACGCCATGGCCGAGTCACCTAGCC 2078
GCGGCACATGTCCTCCATGGAGCACACGGAGGAGGGGCTTCGGGGAGCGACTTGCCGACGCCATGGCCGAGTCACCTAGCC
* * * * *

211 290
2079 GGCAGCTCOTGGGATCCGGAAACAGAACTTCAGCGAGAGGGAGCATCGAGACTCTGAGTAACAGCTCAGGCTCCACCAGC 2124
GG-ACGTCTGGGATCCGGACAGGAAAAAAAAAAAAAAAAAAGG SEQ ID NO:317
* * * * *

291 370
2125 GGCAGCATACCAAGAACTTTGATGGCTACCGATCTCCGCTGCCCAATGAGAGCCAGCCCTCAGCCTCTTCCCGAC 2124

371 412
2125 TGGCTTCCCGTAGGTACCAGCAACCTGCTTCTGACTGGCCAG 2124 SEQ ID NO:315

FIGURE 1B

```

1 1 CGAGAACTCTGAATACAGCTCAGGCTCCACCAGCGGACGATACCAAGAACTTTGATGGCTACCGATCTCCGCTGCCC 0
1 1 60
1 1 GGCAGCCCCCTCAGGCTCTTCCGAACTGCTTCGCGGTAGGTACCAGCAACCTGCTTCTGACTGCCAGGCC 70
81 1 ACCAATGAGAGCCAGCCGCTCAGGCTCTTCCGAACTGCTTCGCGGTAGGTACCAGCAACCTGCTTCTGACTGCCAGGCC 160
1 1 150
71 1 CCTCCCTGCTGGAGGAGGGGAGAAAGCCCCGCTCTGGTCTTACGCTTCAGTCTCTGCTCTTCTTCAACCAACCTGCT 150
161 1 CCTCCCTGCTGGAGGAGGGGAGAAAGCCCCGCTCTGGTCTTACGCTTCAGTCTCTGCTCTTCTTCAACCAACCTGCT 160
1 1 230
151 1 CCAAGCTTAGTGACAGCAGCCGCGCCATCTTACCTGGATGGAGAGAGACCCCTTCTCCAGGCACTCAGCGCACTTCCCT 230
241 1 CCAAGCTTAGTGACAGCAGCCGCGCCATCTTACCTGGATGGAGAGAGACCCCTTCTCCAGGCACTCAGCGCACTTCCCT 320
1 1 300
131 1 CTGCCACACCTGTGCGGTGGAGGCTGTGCCAGGAGAGACTGTAGAAGCTCGGTCCCTGTGTATGTTTGA SEQ ID NO: 318 300
121 1 CTGCCACACCTGTGCGGTGGAGGCTGTGCCAGGAGAGACTGTAGAAGCTCGGTCCCTGTGTATGTTTGCATATGACATCC 400
1 1 480
301 1 TGCATTTGGATCCGCTTTGTATTTTTTAACCATACCCGCGGTGCGGCGCGGTGGGGGAGGCTGGACAGTGACCATCT 300
401 1 480
301 1 GGGGCGCTGAGTGGGGACAGTGTGATCTGCCACCTGCGCCATTTTGACCCGTGAGTGGACAGTCAAGCCCTCAGCTCATGT 300
481 1 560
301 1 CTGGCTGTGACACACACTGCCCCAGCTTCCCTTGGTCAAGCCCCACTCCAGCAGGGGTGAACGGAGGCCAGAGACTA 300
561 1 640
301 1 GGGAGGAGGAGGAGGAGGACATGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
641 1 720
301 1 CTTTAAGCCGCTTTACCCCTGGTCTCTGACTGATCATGAAGGAACCGTGGTTACTGAGGCCCTGTGAAAAGTGCACGT 300
721 1 800
301 1 CTTGTCCAATAAATCAGCTGCGTGTGGAAAAAAGGATCTTTAATTAAGCGGCGGACCTTATTT 300
801 1 880
301 1 CCCTTTAGTGAGGGTAAATTTAGCTTGGCACTGCGCTCTCTTTTACAACCTCGTGACTGGTAAACCTTGGCGTTACCCA 300
881 1 960
301 1 ACTTAATCGGCTTGCAGCACATCCCCCTTTCGCGAGCTGGCGTAATAGCGAAGAGGCGCGCACCGATCGCCCTTCCCAAC 300
961 1 1040
301 1 AGTTGCGCAGCCTGAATGCGGAATGGGACGCGCCCTGTAGCGGCGGCTTAAGCGCGCGGCTGTGGTGGTTACCGCGAG 300
1041 1 1120
301 1 CBTGACCGCTACACTTGGCAGCGCCCTAGCGCGCGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
1121 1 1200
301 1 TTCCCGGCTCAGCTCTAATCGGGGCTCCCCCTTAGGGTTCCCGATTTAGTCTTACCGGCACTTGAACCCCAAA 300
1201 1 1280
301 1 AACTTGATTAGCGTGAATGCTTCACTAGTGGGCAATGCGGCTGATAAGACGGTTTTTCCCGCTTTGACGTTGAGTCCAC 300
1281 1 1360
301 1 GTTCT 1365 SEQ ID NO: 318

```

FIGURE 2

SEQUENCE LISTING

<110> Williams, Lewis T.
Escobedo, Jaime
Innis, Michael A.
Garcia, Pablo Dominiguez
Sudduth-Klinger, Julie
Reinhard, Christoph
He, Zhijun
Randazzo, Filippo
Kennedy, Giulia C.
Pot, David
Kassam, Altaf
Lamson, George
Drmanac, Radoje
Crkvenjakov, Radomir
Dickson, Mark
Drmanac, Snezana
Labat, Ivan
Leshkowitz, Dena
Kita, David
Garcia, Veronica
Jones, Lee William
Stache-Crain, Birgit

<120> Human Genes and Gene Expression Products
XVI

<130> PP-01625.001PC

<150> 60/192,583
<151> 2000-03-28

<160> 324

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 214
<212> DNA
<213> Homo sapiens

<400> 1
ttagtactgc atatgtaaat actacctttt caatgagcta tataaacaat gatagcacat 60
ccttcctttt actatgtctc acctccttta ggagagaact tccttaagta agtgctaaac 120
atacatatac ggaacttgaa agcttttggtt agccttgctt taggtaatca gactagttta 180
cactgttttc agggagtagt tgaattacta taag 214

<210> 2
<211> 353
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(353)
<223> n = A,T,C or G

<400> 2
ggcacgagga gagaactaga aaatatgtat attggatata ctatgtgcc a ggcacgattc 60
caagcccttg atacattctc tannnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
nnnnnnnnnn nnnnnnnnnn nnaaagtgga aactgggat ttgaacaagg ttttggttg 180
gcattctttt ctatgggagc tcagaaatat ctgttgtcta gccctttctc agcctcccaa 240

cctttctcggt	tccttaccta	tgtcacagct	gacttttgagc	taaagtcac	tcggggcagc	300
taggtgccta	tgtgagctgg	cgttcatttc	tcactgtttc	tccttcctaaa	tac	353

<210> 3
 <211> 399
 <212> DNA
 <213> Homo sapiens

<400> 3						
ggcagcagcc	caccaagagc	tgcataagagc	acgttttagct	agagtaggag	tttgcaagtgc	60
tcataatggga	aatgctgctg	ctatactttt	aggaattttct	gagtgcaatt	tagaaacatc	120
tagcacactt	gaaacactgc	gtatcatttt	cctcactcat	gaatatagtc	atcagaattc	180
ataaatagtt	tacctgagcc	ctttaacaac	ctcaaataagg	ccatatttct	ctctctgggt	240
gatggcatgg	accctacagg	aaaaaccaca	ccttaccgct	tctgaccagc	atcactacaa	300
aaaggagtg	tgaagccaat	caccatgtaa	gcaagataaa	agcaaagggg	gtcttgccctg	360
cccctctctg	ttccatacat	tcttaccagg	cactgagag			399

<210> 4
 <211> 389
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(389)
 <223> n = A,T,C or G

<400> 4						
ggcagcagga	gaggggtggtg	ggtccctgag	ttgggtggaaa	gggatagagn	nnnnnnnnnn	60
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnngagtagag	gatgcctggt	atgaggcaat	300
atttgggata	gggaagggaa	gcttggggatt	ttagctacgt	agagacactt	gaaaattgga	360
gggaggaaag	gagtggtggtg	ctttgggagn				389

<210> 5
 <211> 279
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(279)
 <223> n = A,T,C or G

<400> 5						
cctctcccct	tggaaaccaa	agaggaacgg	ggccgaactt	tataaacttt	aggcaagggc	60
aaagggcgtg	nnnnnnnnnn	nnggggccaa	ggggcatttc	ccaagcgatt	aaaatttggtg	120
aacctttggt	tacaaaaatt	gcggggaaaa	tttatttcgg	gagcaatttt	cccttttaaaa	180
atttgagaat	tcttaccg	agagtgtgac	ataatttaag	gcgcctctgc	ccaaagaggc	240
catgtgcgtg	aggggaatac	cgcgtttaat	tatcacaaa			279

<210> 6
 <211> 388
 <212> DNA
 <213> Homo sapiens

<400> 6						
ggcagcagggc	agaggcctcc	ctgcaactggt	cctggcctca	ctcttttccc	tgacccttgg	60
ggcccagggc	catggaggga	cccttaggag	ttcaatgaga	gagaccatga	ggccactggg	120
ctttcccctt	cccaggcctc	ctgggtgcca	cccccttacg	ttattcttgg	gcctctaata	180
agtgtccac	aggtgcctgg	ccaggcccac	ctgctgcaga	tgtggtctgt	gtgtgtgcat	240

gtgtgggtgt	gtgtgggcac	aggtgtgagt	gtgtgagcaa	cagtacccca	ttccagtcgt	300
ttcctgctgt	gactaagtca	gcaacacagt	tcctctgaca	tgggccttgg	ctgtgcttct	360
ttgggggtga	agagattgcg	gaggaagt				388

<210> 7
 <211> 410
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(410)
 <223> n = A,T,C or G

<400> 7						
ggcacgaggg	gaagtcgcgc	atgcgcgagt	gtacgcgttg	ccggcgaaga	ggggagcctg	60
acgactcgga	aatttgaata	ccacagtagc	atggagtgtg	acctcatgga	gactgacatc	120
ttggagtcgt	tgggaagatct	aggttacaag	ggcccattgt	tgggaagatgg	agcgctctct	180
caggcagtcct	ctgctggagc	cagttcccc	gagtttacca	aactctgtgc	ttggctgggtg	240
tctgaattaa	gagtgcctctg	taaactagag	gaaaacgtgc	aagcaactaa	cagtcgcgagt	300
gaagctgaag	aattccagct	tgaggtgagt	gggctactag	gggagatgaa	ctgcccgatat	360
ctttcactga	catctgnnga	tgtgaccaag	cgccttctca	ttcagaaaaa		410

<210> 8
 <211> 229
 <212> DNA
 <213> Homo sapiens

<400> 8						
ctaacaaaaa	acactaaaaa	aaaataaaaag	aaattaattg	aaactgacct	aactcgtggc	60
agggggaact	cggctataag	acccacaaac	cctgctgact	cataacaaac	tgagttgtaa	120
gacattcatc	gccgcgatat	ccttgagtaa	agaatgaact	ctggaagccc	acccacggac	180
aatgcacctt	cacaaagatt	ctgcactaat	ctgagtgaag	gtcctttggt		229

<210> 9
 <211> 380
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(380)
 <223> n = A,T,C or G

<400> 9						
ggcacgagag	tagttgggaa	atcttttata	aatccaccta	ttactaccta	ttggtagggg	60
agattaaatt	tctacaggta	tggagagtcg	gcttgactac	actgtgtgga	gcaagtttta	120
aagaagcaaa	ggtatagcag	ttccaagtan	nnnnnnnnnn	nnnnnagacc	aaactctaga	180
tcttgcccaa	aatggacggc	cgcggcattt	aaatgaagaa	agattttattt	ttcctttttt	240
cttttaagaa	aaattttttt	aaaaaatttt	gattnnnnnn	nnnnnnnnnn	nnnnnnnnnn	300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	360
nnnnnnnnnn	nnnnnnnnnn					380

<210> 10
 <211> 317
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(317)
 <223> n = A,T,C or G

```

<400> 10
cacacacaca catccactct ctcttttttg tctctttctca cacacacata tctcccttac      60
tcacacactc tctctcacac acccctctt tcttttcccc cgcactttct ttctctcaag      120
cgcgcgcgca ctcaactctct tttttcttct ctctctcact ctctctctcc gcgcgctctc      180
tcacacgctt tatactctctc tctctgaggg acttctctct cctctcactc ttattttttt      240
gttggtgttt atagcgtctc tctcttccct nnnnnnnnn ntctatatat acagagagag      300
atctctctgc tctctcc                                     317

```

```

<210> 11
<211> 391
<212> DNA
<213> Homo sapiens

```

```

<400> 11
ggcacgagag aattagctga aaccaccaa gagctgcata gagcacgttt agctagagta      60
ggagtttgca gtgctcatat gggaaatgct gctgctatac ttttaggaat ttctgagtgc      120
aatttagaaa catctagcac acttgaaaca ctgcgtatca ttttcctcac tcatgaatat      180
agtcatacaga attcataaat agtttacctg agccctttaa caacctcaaa taggccatat      240
ttctctctct ggttgatggc atggacccta caggaaaaac cacaccttac cgcttctgac      300
cagcatcact acaaaaagga gtgctgaagc caatcaccat gtaagcaaga taaaagcaaa      360
gggggtcttg cctgcccatc tctgttccat a                                     391

```

```

<210> 12
<211> 280
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(280)
<223> n = A,T,C or G

```

```

<400> 12
tgtgcgcgcc cccccggggc gctctctctc tacactcgtg cgctcccccc tctgtctgtc      60
tctctctcta gagtcacggt ctcttacacg gcgcgcacat gcgaggggca ctnnnnnnnn      120
nnngctcnnn nnnnnnnnnn nnnnnnnnnn cgnnnnnnnn nnnnnnnntcc cttgtatact      180
ctctgtgtgc gcggggacan nnnnnnnnnn nngtgcgcgc gcgagagcgc gcgcgccaca      240
caagagagag cgcgccctnn nnnnnnnnnn naccgcgaac                                     280

```

```

<210> 13
<211> 311
<212> DNA
<213> Homo sapiens

```

```

<400> 13
cgcttttttg ggaacccaaa cctttttttg ctggccggaa aaaattttcca cgggaagggt      60
aaaggggttt attaatTTTT ttggcaaaac aggggttaag aaaccttccc tcccggccta      120
agggtggggt aggcTTTtga aaggctaaaa gggggaaatt tctggccctt gttccaaggg      180
aaacatgggc tagggggaaa cccacccctc tcagggccct ttaaaagggc ccccaaaaaa      240
agaaccctt tattaagggt taaaaaagg taaaaaagg gggaaacctc tgggccaagg      300
caaatttttg t                                     311

```

```

<210> 14
<211> 387
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(387)
<223> n = A,T,C or G

```

```

<400> 14

```

```

ggcacgaggt cttttctgcc cacatctcac acaattgagg tgtctgaaca agcttgggga      60
gggtctataa ggggtaggct cnnnnnnnnn nnnccattt ggaaagggcg ttttgccaac      120
ccaagggctt ttttaagccg atttttnnnn nnnnnnccgg acttggtaat tggccttttg      180
ctttttaaag cccaaaaaat aataattaag gggcccaaaa taaggaaggg caaaaaaagc      240
ctttactccc cctgcctttc aaaaagaaaa ggaaaaaccg gccccccctt aataattggc      300
accctaataa aaaggggttt taaaaaaagc caaaaacaaa agggcctgga aaaaaatttt      360
gacttttttt aaccgcgaac ctgggaa                                387

```

```

<210> 15
<211> 273
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(273)
<223> n = A,T,C or G

```

```

<400> 15
ctgtctctct ctctccccc ctctccctcc cgcgcgcgca cgctctttca tctctctctc      60
tacagacagg ggggggtgtt ctctctccct ctcgagagg accgcttttt tttctcccc      120
ctctctcaca ctcggggtgt gcgcgctccc tttgggggct tttctatagg gcgcgctcta      180
aagaaagccc gcctttctcc tctgggtgcc tctccacaca cccgggtttt ctcccccgct      240
gtttttgaag aaactcctcc tggctctcct atn                                273

```

```

<210> 16
<211> 283
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(283)
<223> n = A,T,C or G

```

```

<400> 16
ctctctctct ggcccccccc ctctctttac acacactttc tctcctctct ctcgctctct      60
cttttttttt ctctctcccc tcgctctctc tgtgtgtctc tatctcgtgt ctctctctgc      120
gtgtccctca cacacactcg cgcgagagat ctctctctat atctctcctt tgtctgtgtc      180
tctctctcgc gcgccacac atctatatat ttttgcgcg acacgcgaga gtgtgtccct      240
ctctctctct gcacnnnnnn nnnnnnnnnn cacaccctcc ccc                                283

```

```

<210> 17
<211> 392
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(392)
<223> n = A,T,C or G

```

```

<400> 17
ggcacgagat gactccttcc ttaaaatcca gctcaaatct cccccctttt ggtggctttc      60
tctgacactc catcataaag ctaattgttt aagtatgata cagtggcaca gtttattcct      120
acttcataac ttttatctca ctatgttgta agatattagg tatgtttctt ctactaccag      180
taattttcaa agagttaagg aagaaggata gaagacagca gtatagggtga atgtgtgcat      240
gtgttnnnnn nnnnnnnngc catattggcc aaaatttttg gactggctgg taaaacaaag      300
gcttttcaaa ttttcaaata ctttataaaa aaacctggaa attgttttgt nnnnnnnnnn      360
cgcccaaaaa aaaatttttg gcctgggggg ga                                392

```

```

<210> 18
<211> 385

```

<212> DNA

<213> Homo sapiens

<400> 18

ggcacgagggc	agaggcctcc	ctgcactggg	cctggcctca	ctcttttccc	tgacccttgg	60
ggcccagggc	catggaggga	cccttaggag	ttcaatgaga	gagaccatga	ggccactggg	120
ctttcccctt	cccaggcctc	ctgggtgcca	cccccttacg	ttattcttgg	gcctctaata	180
agtgtccac	aggtgcctgg	ccaggccac	ctgctgcaga	tgtggtctgt	gtgtgtgcat	240
gtgtgggtgt	gtgtgggcac	aggcgtgagt	gtgtgagcaa	cagtaccca	ttccagtcgt	300
ttcctgctgt	gactaagtca	gcaacacagt	tcctctgaca	tgggccttgg	ctgtgcttct	360
ttgggggtga	agagattgct	gaggc				385

<210> 19

<211> 383

<212> DNA

<213> Homo sapiens

<400> 19

gaaggcttgc	ggagagaaaa	ccctggagcc	atcttcatag	gaagaggaaa	ggaaactgta	60
tgacaggaga	atgaatcaag	tttggggctc	aaggtgccgg	ccactgggaa	aaacagctgc	120
cccagagttgc	aaaactctgg	gtcctatatg	tataaactat	gccctgagga	aggaatctca	180
ggcgtatctt	aggagaaaaat	gttctagctt	gggaaacaaa	cacaacagga	ccgtgaatcc	240
aaatatattca	agtgggttta	gaggactgga	gttctaaacg	ctgctttttac	tgtaagtgat	300
cacgccccgg	aatgtgctga	agaaaggaaa	atgagccagt	atcggcgagg	actatgggca	360
aggaaaacga	gagtgtgcga	tgt				383

<210> 20

<211> 313

<212> DNA

<213> Homo sapiens

<400> 20

ctctcccccg	cgctcttgag	atatgcgcgc	cccttttttc	ttctacacgg	ggggggggcgc	60
gcctctttttt	ctgcgcgcgc	cccctctctc	tcttttgtgc	gcacgcgcgc	gcgcgggggg	120
gttctttttt	tgtgcggaga	gagagtctgt	ctcaggggtt	tttttgtttt	ctttcacgac	180
acacactttc	tcccctgtgc	atgtgttttg	atgctctctc	gagatatgtc	tctctctctc	240
tgtgtgtgtg	tgttgtgcgc	ccccctggg	gagagcgctc	ttctctctct	cctcatatag	300
cgcgcgcgcg	cga					313

<210> 21

<211> 396

<212> DNA

<213> Homo sapiens

<400> 21

ggcacgaggg	gaccccccttc	acctctgtct	agagagctgg	gtagatcaga	aacttggtga	60
cacctggcta	gcacagagca	ggctcacttg	tcttgggtccc	actaccaga	ttcctgcaga	120
cattgcaaac	caaataaagg	ttgttgaatg	acccctgtcc	ccagccactt	gttttgttat	180
catctgctct	gcagtggaaat	gcctgtgtgt	ttgagttcac	tctgcatctg	tatatattgag	240
tatagaaacc	gagtcaagtg	atcatgtgca	tccagacaca	ctgtgtcacc	tgagccacag	300
agcaaatac	cttaacgatc	tggaatgaaa	ctgtgaccag	tgccgccttg	ggtgggttctg	360
gagagactgc	cgtcttcttg	tttggccata	ggtgcg			396

<210> 22

<211> 310

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(310)

<223> n = A,T,C or G

```

<400> 22
tgaatatcag ggccctgaac atctctcacg cccgtcttctt aaaagagaag aaaaaaacgc      60
gcgcggggctt tttctctctc tcagaggggt gaaacacaca atatctcggg gggccggggg      120
agagcccgcct ctctctgcct gtaaaacaca cagaagtgcg ctcacgccct gcgcggggagc      180
ccacagactt ttttttaaaa caaaaagtat attgggggtgt gttttaatct ccctctccgc      240
tcctagaggg ggggcggnnnn nnnnnnnnnn ntttttaaat agggggggccc gagtctcacc      300
caatagaagg                                     310

```

```

<210> 23
<211> 375
<212> DNA
<213> Homo sapiens

```

```

<400> 23
ggcacgagcc ggccaagagg ggagcctgac gactcggaaa tttgaatacc acagtagcat      60
ggagtgtgac ctcatggaga ctgacatctt ggagtcgttg gaagatctag gttacaaggg      120
cccattgttg gaagatggag cgctctctca ggcagtctct gctggagcca gttccccga      180
gtttacaaaa ctctgtgctt ggctggtgtc tgaattaaga gtgctctgta aactagagga      240
aaacgtgcaa gcaactaaca gtccgagtga agctgaagaa ttccagcttg aggtgagtgg      300
gctactaggg gagatgaact gcccgatatc ttcactgaca tctgggggatg tgaccaagcg      360
ccttctcatt cagaa                                     375

```

```

<210> 24
<211> 477
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(477)
<223> n = A,T,C or G

```

```

<400> 24
gctccttctt cttnntgttg atcccatcga tccgaattcg gcacgagagc acctctgtgc      60
ctctctgaga gcactcacag ccaaaagtac acagctgccc ccaggctgag agtgcttgat      120
acacccttga atccctctct atatgatgcc ccagcccagg agagataaaa gcatcagcac      180
catgagattc acctgcctct ggtcgtnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnact cttagacagc aaaaatgctt      360
tctcccagtc ttgttccctt gttctcagtt cccaccctgc ctggataact actgttcttg      420
gtttnnnnnn nnnnnnnnnn nnnnnnnnag tctcgtacca gattcataaa tcagccg      477

```

```

<210> 25
<211> 265
<212> DNA
<213> Homo sapiens

```

```

<400> 25
cgcgcggggg ggacccctct ctctctctct gttgcgcgcg ctctctcacc ccgtgtgtcg      60
ccccgatata tgtcagagag accccctatt tttttctccc gcccacaca catctatgtg      120
taaaatgtgc gtgtctgtcg cgcacacca cacactctcc ccgggggggt tataaaatac      180
tcgcgcgcta tattttcgcc cccctttttg tgtgtgggcg ccacaaaaac accacacgct      240
ctccccctg tctctcgcgg gtgtt                                     265

```

```

<210> 26
<211> 388
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(388)
<223> n = A,T,C or G

```

<400> 26

ggcacgaggg	aggctcttttg	ttatagatgc	ttttgcccc	ttaatacagc	aatgagagca	60
ctgaccgaag	aggcagccgt	gactgtaaca	cctccaatca	cagcccagca	agctgacaac	120
atagaaggac	ccatagcctt	gaagttctca	cacctttgcc	tggaagatca	taacagttac	180
tgcatacaacg	gtgcttgtgc	attccaccat	gagctagaga	aagccatctg	caggtgtcta	240
aaattgaaat	cgccttacaa	tgtctgttct	ggagaaagac	gaccactgtg	aggcctttgt	300
gaagaatttt	catcaaggca	tctgtagaga	tcagttagcc	caaaattaaa	gttttcagat	360
gaaacaacaa	aacttgtcaa	gctgactn				388

<210> 27

<211> 431

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(431)

<223> n = A,T,C or G

<400> 27

ggcacgagag	aggggctact	ttagatgcaa	aggggacaat	tagaaggcta	ctgaggtaat	60
ccggacaaaa	agttgtaaat	aaatcacggt	ggcagtatgg	tgaatagtgg	aaggggtgta	120
tttgaagaaa	ctggggaggg	cgtgggagag	gctggctagt	gagaaatggg	ccgaagggtga	180
aagcagctta	ggggctggtt	tccagttttc	tggcactgca	gactgggtag	tgggaggtgg	240
ctttctcaag	aggagaggtg	agtgggaagg	agcagggctg	caggggaggt	catggtcttg	300
ggagtgggtgc	tcagtctgac	ttgcacatag	gggagattat	tttagatttc	cgcaagaaaa	360
tgtccagcat	gtagtcatat	caatgnnnnn	nnnnnnnnnn	nnnnnnnnnn	nntgagattt	420
acccaaaaag	a					431

<210> 28

<211> 389

<212> DNA

<213> Homo sapiens

<400> 28

ggcacgagcc	accccccaaga	gtgtggccat	ctggggccgt	gtggtatttg	ccactcagga	60
gacatgtccc	tatgacatag	cagtgggtgag	cctggaggag	gacctggatg	atgtcccat	120
ccctgtgccc	gctgagcact	tccatgaagg	cgaggctgtg	agtgtgggtg	gctttggcgt	180
ctttggccag	tcttgcgggc	cctcgggtgac	ctcaggcatc	ctttccgctg	tgggtgcaggt	240
gaatggcacg	cccgtaatgc	tgcagaccac	gtgtgctgtg	cacagcggct	ccagtggggg	300
accctcttc	tccaaccact	caggaaacct	ccttggcata	atcaccagca	acaccoggga	360
caataatacg	ggggccacct	acccccacc				389

<210> 29

<211> 431

<212> DNA

<213> Homo sapiens

<400> 29

ggacgaggct	ccagcgcact	tttccaacac	atcactgcat	tatttgaatg	caccatggca	60
gctattgtca	ccttacttgg	gagtgatcca	gttggagctc	tttatattcg	gacatgtcga	120
gtattgatgc	tttctgactg	ggacacgatg	ctttacaacc	caaggccaga	ttacgggtacc	180
acagtgcact	gtactcatga	agccggctac	ccactatata	ccatcgattt	tatctattac	240
gcattctgct	tggtattaat	gatgctgctc	cgacctcttc	tggtgaagaa	gaatgcatgt	300
gggttaggga	aatctgatcg	atataaaagt	atztatgctg	cactttactt	cttcccaaat	360
gtaaccgtgc	ttcaggcagc	tgtgggaggc	cttttatatt	acgccttccc	atacattata	420
ttagtgttat	c					431

<210> 30

<211> 393

<212> DNA

<213> Homo sapiens


```

<400> 30
ggcacgagac tacacccgct tcgatgactg gtacctgtgg gttcagatgt acaagggggac      60
tgtgtccatg ccagtccttcc agtccttgga ggcctactgg cctgggtcttc agagcctcat      120
tgagacattt gacaatgcca tgaggacctt cctcaactac tacactgtat ggaagcagtt      180
tggtgggctc ccggaattct acaacattcc tcagggatac acagtggaga agcgagaggg      240
ctaccacatt cggccagaac ttattgaaag cgcaatgtac ctctaccgtg ccacggggga      300
tcccaccctc ctagaactcg gaagagatgc tgtggaatcc attgaaaaaa tcagcaaggt      360
ggagtgcgga tttgcaacaa tcaaagatct gcg                                     393

```

```

<210> 31
<211> 459
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(459)
<223> n = A,T,C or G

```

```

<400> 31
gcaatcgcat tgtctttttg aggatnnnat naatgtcaat tcggcacgag ctttgtggat      60
gtttccagct gccatcgta ccttctgtgc tgctccctgg accagcttca ggacttgaag      120
gccctcgtgg ctgagatcat cacacatttg caggggctgc agagggactt atctctagca      180
gtctcctaca gcaggctcca ttcctcagac tggaatctgt gtactgtatt tgggatcctc      240
ctgggctatc ctgttcccta tacctttcac ctgaaccagg gagatgacaa ctgcttagct      300
ctgactccac tacgagtatt cactgcccgg atctcatggt tgctaggtca accccaatc      360
ctgctctatt cttttagtgt cccagagagt ttgttccac gcctgagga cattctgaac      420
acctgggaga aagacctcag aacccgattt atgactcac                                     459

```

```

<210> 32
<211> 445
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(445)
<223> n = A,T,C or G

```

```

<400> 32
ggcacgagat ggagagcacc tctgtgcctc tctgagagca ctacacagcca aaagtacaca      60
gctgccccca ggctgagagt gcttgataca cccttgaatc ccctcttata tgatgccccca      120
gccagagaga gataaaagca tcagaccatc gagattcacc tgcctctggt cgtnnnnnnnn      180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      240
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      300
nnnnactctt agacagcaaa aatgctnttc tccagtcctt gttccttggt ctcaaggtcc      360
acccttgctg gataactact ggtcttgggt tcctggggta aagatggaac ttgagtaagc      420
tcgacccaaa tccaaaatca atccg                                     445

```

```

<210> 33
<211> 429
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(429)
<223> n = A,T,C or G

```

```

<400> 33
ggcacgagcg cctgccctgc atcagggaga catgtcagct gaggagtaat tgaccagatt      60
tctgcttttag aaatatggca gtggaggcag gagatggcat ctgaggccca ggctggggag      120
aagggtgctg ggatgagaac ctggagttca gaccaggga gggatgagag cctaagaaga      180

```

ggagctctca	ccctgagaca	ggctggtgca	ggagtctgct	cgatccaggc	ctgggtccct	240
ggttccctct	gagcttgga	ggactatgtg	agacagaaca	ggaccagggg	cctgcattcc	300
cccttgatt	attcatcn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnna	420
tgatggccc						429

<210> 34
 <211> 439
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(439)
 <223> n = A,T,C or G

<400> 34						
gttctgtggg	aatagagggt	ccctggtgac	agggcagggc	tagatctgga	gcctgcactt	60
ggcctgtgac	atactgtctt	gtttctgaga	atcctcccct	acttctctag	ttaatctcca	120
gagacttctg	tgactactta	atcacaaagg	aaattttcag	gaatattatc	aaatactatt	180
ttagaaaaaa	aaagagaagg	gatttgaatg	ttttcagttc	agtttagnta	tcnnnnnnnn	240
nnnnnnnccc	caaacttcaa	aatggaggcc	ccccctcct	ttaaccccc	taaaaaaaat	300
tctgatgttt	gagggttgg	tgccaattaa	ccaaaccccc	aaaaaaaaag	ggggttaaac	360
cccattggaa	agttttccta	attttggggg	gtgccctttg	aggtggaccc	ggttccctgc	420
cctgggaaag	gccccaaag					439

<210> 35
 <211> 440
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(440)
 <223> n = A,T,C or G

<400> 35						
ggcacgaggt	gaagtcctgg	ttccagactc	ccctttttgc	cgggacatga	tggatctgtc	60
agctggtgcc	tatagtccta	gagagctaga	gatggaggga	aattcagatc	atctaaaccc	120
ttcagccctt	cactggacag	aagaggaaac	tgaggctcca	tctgcatgac	gttcccagag	180
tcacggcaca	aattcatgga	agaagcagca	ggaaactcag	ttctccagtc	tgggtccaat	240
gtgtgtttta	gaaatatctc	cacagggtta	atgactcaat	ttttcatgca	tgattgctag	300
taatgacaat	catgttatgt	ttgtttctgt	agctttggaa	atcactcctt	ccacttgagt	360
ttcaggtccc	aactgtccac	acctgcagga	gtgaggtttt	gctgagactg	ataaggcact	420
cacattntgt	gggagttgaa					440

<210> 36
 <211> 423
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(423)
 <223> n = A,T,C or G

<400> 36						
acgagcgenn	nnctgcatc	aggagacat	gtcagctgag	gagtaattga	ccagattttct	60
gctttagaaa	tatggcagtg	gaggcaggag	atggcatctg	aggccaggc	tggggagaag	120
ggtgctggga	tgagaacctg	gagttcagac	cagggaaggg	atgagagcct	aagaagagga	180
gctctcacc	tgagacaggc	tggtgcagga	gtctgctcga	tccaggcctg	ggtccctggt	240
tccctctgag	cttgggagga	ctatgtgaga	cagaacagga	ccaggggcct	gcattccccc	300
ttgtattatt	catcnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	360

nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnatga 420
tgg 423

<210> 37
<211> 424
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

<400> 37
ggcttgtaga nctcggaggt tngcaagaat cgcattcggc acgagctggg acacagtngg 60
ctctcttata tttgttgctg gaataaatga atgaactaag gcagtcttgt agggatttac 120
tggttaaccac catgggaaaa ttaaataaat gcggggaagg aaaacgttct aaaattagaa 180
gactactttc tactctcagc ttctgattcc ctctgagcta agaaccagac agccttaggc 240
tggttaactcc tataagctgg tcctcctccc atgctgaccc catctttact gtacaattca 300
cttttcatgg actgaaggca ccaccaagat agatccagga gtgacaactc cagtgtagggt 360
gtccactggt cccttaatct ctgtcctgct ccaagtataa ataaatcggg gccatttcct 420
taga 424

<210> 38
<211> 434
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(434)
<223> n = A,T,C or G

<400> 38
ggcacgaggt acacagctgc ccccaggctg agagtgtctg atacaccctt gaatcccctc 60
ttatatgatg cccagccca ggagagataa aagcatcagc accatgagat tcacctgcct 120
ctggctgtnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
nnnnnnnnnn nnnnnnnnna ctcttagaca gcaaaaatgc tttctcccag tcttgttccc 300
ttgttctcag ttcccaccct gcctggataa ctactgttct tggtttnnnn nnnnnnnnnn 360
nnnnnnnnnn agtctcgtac cagattcaaa aatcagtcaa ctacttcaaa aacaatgaca 420
tgctggctac ttaa 434

<210> 39
<211> 428
<212> DNA
<213> Homo sapiens

<400> 39
ggcacgagct ttgtggatgt ttccagctgc cagcgtcacc cttctgtctg ctccctggac 60
cagcttcagg acttgaaggc cctcgtggct gagatcatca cacatttgca ggggctgcag 120
agggacttat ctctagcagt ctccacagc aggtccatt cctcagactg gaatctgtgt 180
actgtatttg ggatcctcct gggctatcct gttccctata cctttcacct gaaccaggga 240
gatgacaact gcttagctct gactccacta cgagtattca ctgcccggat ctcatgggtg 300
ctaggtcaac ccccaatcct gctctattct tttagtgtcc cagagagttt gttcccaggc 360
ctgagggaca ttctaaacac ctgggagaag gacctcagaa cccgatttag gactcagaat 420
gactttgc 428

<210> 40
<211> 429
<212> DNA
<213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(429)
 <223> n = A,T,C or G

<400> 40
 ggcacgagtg gagagcacct ctgtgcctct ctgagagcac tcacagccaa aagtacacag 60
 ctgccccag gctgagagtg cttgatacac ccttgaatcc cctcttatat gatgccccag 120
 cccaggagag ataaaaagcat cagcaccatg agattcacct gcctctggtc gttagggaac 180
 aatggaggcc tgcgatttgg agttaaaactc tcagtgatct ctgtgttgac aacaccaaag 240
 ctagaggaat ccagtaggat gtgggcatgg ttttcccga aggctgactg agcagttctg 300
 caaatgtttg caagtacagg gcagaatttc atccagcctc agaaccctga gccaaagactc 360
 agcatcagca aagccaaaag tttcattttc ttgactgtgg gagtgcctagt cccaaccttt 420
 agatggccn 429

<210> 41
 <211> 430
 <212> DNA
 <213> Homo sapiens

<400> 41
 actctgcaaa cagctacttg tgctgattgc aggagaccca taaattcgaa cgaggaacaa 60
 ccgagacctg aaggggctga cgaacgcgat ttctgataag tatgggggtcc ctgaagagaa 120
 catttaccaa gcctacaata aatgcacgag aggaatctta tgcaacatgg acaacaacat 180
 cattcagcat tacagcaacc acgtcgcctt cctgctggac atggcggagc tggacggcaa 240
 aattcagatc atccttaagg agctggaagg cctctcgagc atacaaaccc tcacgacctg 300
 catggggcca gcagggacgt ggccccacgc cacacacaac ctctccacat gcctcaacgc 360
 tgttacttga atgccttccc tgagggaaga ggcccttgag tcacagaccc acagacgtca 420
 ggaccatggg 430

<210> 42
 <211> 437
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(437)
 <223> n = A,T,C or G

<400> 42
 ggcacgaggc gccctctgcc cccctcagag ggtctctcct ctcgaccccc aaattccccc 60
 agcatctcaa tcccttgcat ggggagcaag gcctcgagcc cccatgggtt gggctccccg 120
 ctggtggctt ctccaagact ggagaagcgg ctgggaggcc tggccccaca gcggggcagc 180
 aggatctctg tgcgtgcagc cagcccagtg tctgatgtca gctatatgtt tggaagcagc 240
 cagtcacctc tgcactccag caactccagc catcagtcac cttccagatc cttggaaagt 300
 ccagccaact ctctctccag cctccacagc cttggctcag tgtccctgtg tacaagaccc 360
 agtgacttcc aggtcccccag aaacccccacc ctaacccatgg gccaaaccag aacacccccac 420
 tctccaccac tgggcan 437

<210> 43
 <211> 432
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(432)
 <223> n = A,T,C or G

<400> 43
 ggnncagtga ccaccaggac ctggtgtctg tgcacatcta catcaccag ctggctgaga 60
 agttcgacct caggaccact atgctgtaca tctgtgagcg gcacttccag aaggttctga 120

accggagtct	attcacaggc	ctgcgctcca	tcacccactt	tggccgtccc	ccctttgagc	180
ccttcttcaa	ctccctgcag	gaggtccacc	cccaggctcg	gaagatcggg	gtgttttagct	240
gtggccccc	tggcatgacc	aagaatgtgg	aaaaggcctg	tcagctcatc	aacaggcagg	300
accggactca	cttctcccac	cattatgaga	acttctaggc	cccttcccgg	gggttctgcc	360
cactgtccag	ttgagcagag	gtttgagccc	acacctcacc	tctgtttcttc	ctattttctgg	420
ctgcctcagc	cc					432

<210> 44
 <211> 436
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(436)
 <223> n = A,T,C or G

<400> 44						
ggcagcagcc	gaggcgcgcg	tggtccgtgg	ccgcttccag	ggccgcgcgg	cggtgatcaa	60
gcaccgcttc	cccaagggct	accggcaccc	ggcgctggag	gcgcggcctt	gcagacggcg	120
gacggtgcag	gaggcccggg	cgctcctccg	ctgtcgccgc	gctggaatat	ctgccccagt	180
tgtctttttt	gtggactatg	cttccaactg	cttatatatg	gaagaaattg	aaggctcagt	240
gactgttcga	gattatattc	agtccactat	ggagactgaa	aaaactcccc	agggctctct	300
caacttagcc	aagacaattg	ggcaggtttt	ggctcgaatg	cacgatgaag	acctcattca	360
tggtgatctc	accacctcca	acatgctcct	gaaaccccc	cttgaacagc	tgaacattgt	420
gctcatagac	tntggg					436

<210> 45
 <211> 300
 <212> DNA
 <213> Homo sapiens

<400> 45						
tctctctctc	tctctctcac	agacactttt	accccatata	tacacataaa	atgtgtgcgc	60
gagagagaga	gagccctctc	gctctatata	tatccccgcg	ggggggagat	aaaaatatat	120
atccccacac	tttatagggc	gggctcccc	ctctatcctg	tgtgtagaga	gaaatatata	180
tatatctgtg	gggggagaga	gagatctctc	acccccccgc	acacgcgagc	tctttcttaa	240
gatgtgtgag	cgcccccccc	ctgtttttgt	aaaaaagaga	ggggtatata	tattgggggg	300

<210> 46
 <211> 191
 <212> DNA
 <213> Homo sapiens

<400> 46						
caaaacaaaa	ccatgttccc	actggtgatg	cctgtctgac	acgttttggg	atttagtagg	60
aaatgaagg	tcttcaagct	tcgagagAAC	cttcaaaatt	gtcacaattg	ctgaaaacag	120
aatgaatcgg	gaacattatc	tcaatatatt	gcataataga	caacaccaca	gtgttttggg	180
tccctgacct	g					191

<210> 47
 <211> 302
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(302)
 <223> n = A,T,C or G

<400> 47						
gcccgggctg	gtgtgtatgt	gtgtacacgc	ccccgtgggc	tctctgtcgc	atcttgnnnn	60
nnnnnnnnnn	nnnnnnnnnn	nntgtannnn	nnnnnncaca	tagcgcgcgc	gctcgcgcgc	120

acggagctat	agagacacca	ctctctctct	gagatacacg	cgcgcgca	cactctgcgc	180
gcgcgcgctc	ttctttgtct	cgcgcgcgcg	cccgtatgt	ggaggggtata	tgtgggggaa	240
aatagcgagg	tgtgcgcgca	cccgcgcacg	cgcgctctat	atctctatat	cttcagcgcg	300
cg						302

<210> 48

<211> 411

<212> DNA

<213> Homo sapiens

<400> 48

ggcagcaggc	ttgcggggca	ttaggactag	agggttggtg	aaaattcaga	cagaatgtaa	60
cttgacaaag	agaagacagc	aacaactgta	acaattatct	tatgaatatt	tgcgaaactc	120
aaagggatct	gattgggtgac	ctctgggctt	tatcaaatta	acatcacaac	ttctagaaga	180
aagtcaacct	tcatctttta	caatagaaat	catatgtttt	gctaaccat	tcctatttag	240
gctgaaaaca	attaagagtt	atgggtactt	aaaaaaatca	ttatgtttat	aaaattagtg	300
atagaaggag	catagtgttc	tatacagtca	cacacatata	cttccttatt	tcttttatatt	360
aaactttgag	taacatagca	gtctatgttt	gggtcagttt	tccctttttt	g	411

<210> 49

<211> 408

<212> DNA

<213> Homo sapiens

<400> 49

ggcagcaggc	acacaaagcc	aagggcatac	cctatagagt	aaagctgcag	ccaccctgtg	60
tctcatgtgc	agctgaaata	gtgatctgct	tctgtcactg	tcacatagac	agccctgcat	120
gccccctgtc	tcacacagtt	tgtaatgaag	acagctcctt	ctcatctttc	cataagcctg	180
agatacaagt	tcagggactc	agcaatgcac	tttaggactg	agctaggagg	caaatatctg	240
aagcttgcta	tgctgttctt	tccattcctt	ttccctctga	aacacacaaa	ataccaaagg	300
aacttacgca	tcacaccact	gagtcctcta	actaatcata	tgtgctcaga	cacagctcaa	360
gcacaccctt	tagttaagag	agaacctcca	tatacatata	tttttttc		408

<210> 50

<211> 407

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(407)

<223> n = A,T,C or G

<400> 50

agagaaacat	ccactcgaat	tgggcacgag	gacagggcag	ggctagatct	tttttctgca	60
cttggcctgt	gacataactgt	ctgggtgtctg	agaatcctcc	cctacttctc	tagttaatct	120
ccagagactt	ctgtgactac	ttaatcacaa	aggaaatttt	caggaatatt	atcaaatact	180
attttagaaa	aaaaaagaga	agggatttga	atgtttttcag	ttcagtttag	ttatcnnnnn	240
nnnnnnnnnn	nccaaactc	aagtatggag	gccccccctt	ctttaaacc	acaaaaaaa	300
ttttttgggg	ttcaggggtg	gttggccaac	tacccaaacc	cccaaagaaa	atgggggtta	360
accccttga	aaaagttttc	ttactttggg	gggctgcctt	tgagccg		407

<210> 51

<211> 312

<212> DNA

<213> Homo sapiens

<400> 51

ccccgggggc	gctctctttt	tttttcccc	caagtgaag	agccccgcgc	gcgtctctct	60
ctcgcatttt	ttcgacaccc	cccttgtgtg	ggcgggggcg	cgcgctctgtg	tgtgatacac	120
agaatgtgcg	tgggtgtgtc	gagagacact	cttcgcgctt	gtgtgtgaga	cacgagactt	180
tctcttttta	gggggcgggg	ggggagtttt	atgtgtgcca	catgttttct	gtgtataaaa	240
agagcgcaca	gagtgttttt	tatatctgtg	agagagacct	ctctgtatat	atacacgctc	300

agaggggaga gg

312

<210> 52
 <211> 420
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(420)
 <223> n = A,T,C or G

<400> 52
 acgaggggnn nnaagcaccg cgggtacccc atgagggcct acaagctggc caccctggcc 60
 atgacccatc tcaacctgag ctacaatcag gacacacacc ctgccattaa tgatgttttg 120
 tgggcctgtg cgcttagcca ctcccttggt aaaaatgagc ttgcagctat aatacctctg 180
 gtggtcaaga gtgtcaagtg tgcaacggta ctgtcagaca ttttgcgagc atgcactctg 240
 accactcctg gcatgggtggg acttcatggg aggaggaact ctggtaagct catgtcactg 300
 gacaaagccc ccttgaggca actcttgat gccacgatcg gggcctacat caacacaacg 360
 cactcacggc tcacacacat cagtcctcgg cactatagtg agtttataga gttcctcagc 420

<210> 53
 <211> 394
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(394)
 <223> n = A,T,C or G

<400> 53
 ggcacgaggt gtggatgaca gagcgagacc ctgcctcatn nnnnnnnnnn nnnnnnnnnc 60
 ccccccnnnn nnnnnnnnaaa aacccggttg ggccccggtt gttcttttagg gccctaaaaa 120
 ttgccccaaa aaaaattggc cgggccctaa aaaaaccccg gttttttggg gagaattcaa 180
 aaaagggctg gttnnnnnnnn nntttttaaa cttccaaccg gcctcagggg gaaaaaacct 240
 ggaaaactca atggggggttg gaacaaaatc aatatttggt cctaccggaa agcgtaaga 300
 ttttaaacca gtaaaaatgg ccaannnnnn nnnnnnnnnn nnnnaacagg gcccccgagg 360
 taagggctaa aaattttcag atttgaacct tttt 394

<210> 54
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 54
 ggcacgagat tttcttgcca ataagcggac tctgggactc cggctcccta ccccaaactg 60
 aagcgcttcc gtgaacaccc ccgtcctccg tagggggagg ggagcaggcg ggatcctggg 120
 tccctcataa gcacttttgt ttaccgcct gcaacctcac tgtgcccgcc ccgcaccatg 180
 ccctagcccc aggtctagcc gggccattg cagggggcag cacttggggg catctccggc 240
 acttgggtgg gaccaaggag atgccaccat agacctttcc ctgccttct tctccctag 300
 tccgggttcc attcttttca ccagcaccca tcgcccaagg ggtaccgagg gggggcaggg 360
 ggtggtcaat tcaaacccaa cccccgctcg 390

<210> 55
 <211> 280
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(280)
 <223> n = A,T,C or G

```

<400> 55
tctctctctc tctctgcgcc cacacctctc tcannnnnnn nnngcacgtg tatatctnnn      60
nnnnnnnnnn ttttttttag agagacatct cgcgcgtgtc tctctttttc ccgcccgcgc      120
ctctttttctc gcgcgcgcgc gcaccccccc tgtgtggggc gcgcgctctc tttttttttg      180
tgcgcgcgan nnnnnnnnnt ctctctctgt ggcgnnnnnn nnnnnntctc ttattttata      240
ttttgggggg cggggggcct cccctcccc cttgtgtgcct      280

```

```

<210> 56
<211> 398
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(398)
<223> n = A,T,C or G

```

```

<400> 56
ggcacgaggt ccacctcagc tcagcaatct catgccggtt ggcaattagt cagcataagc      60
cgatgcctgc ccatcagttc tttactctga ggtgttagag tggaataaaa atataaatac      120
ttacnnnnnn nnnnnnnnca ataccacaacc ccctcccatt nnnnnnnnnn nnngcccgc      180
cccctaaaat tcatggagag gcctattttc tagccagcca ctatataaac cctgctgggt      240
gggcggnnnn nnnnnnnngt gaagggggga aaaaaagcc tttttttgaa aaaattagtc      300
attttttgct ttttttgac acattttgcy ggacaaagaa ccctgtaaaa cccccctatt      360
cnnnnnnnnn nnnnnnaacc tcaacgaggg gggggcgg      398

```

```

<210> 57
<211> 386
<212> DNA
<213> Homo sapiens

```

```

<400> 57
ggcacgagat tttcttgcca ataagcggac tctgggactc cggctcccta ccccaaactg      60
aagcgcttcc gtgaacaccc ccgtcctccg tagggggagg ggagcaggcg ggatcctggg      120
tccctcataa gcacttttgt tttaccgcct gcaacctcac tgtgcccgcc ccgcaccatg      180
ccctagcccc aggtctagcc gggcccattg cagggggcag cacttggggg catctccggc      240
acttgggtgg gaccaaggag atgccaccat agacctttcc ctgcgccttct tccctccctag      300
tccgggttcc attcttttca ccagcaccca tcgcccaagg ggtaccgagg gggggcaggg      360
gggggtcaag tccaggccca cccccg      386

```

```

<210> 58
<211> 202
<212> DNA
<213> Homo sapiens

```

```

<400> 58
cactttttct atatgaatat cttggccgta tcatagactc aaaaaagaaa ttatgcaagt      60
tctttctgcc cccacctgcg ccagggggaga agtttacctt cgggaactcc agagttaaag      120
cagttgtggg gataattttt tatgctgaac acaccacgat ataaaaaaca acattcacgt      180
gctttatttt tggtatgtgt tt      202

```

```

<210> 59
<211> 394
<212> DNA
<213> Homo sapiens

```

```

<400> 59
ggcacgagtc tgcttctgtc actgtcacat agacagccct gcatgcccc tgtctcacac      60
agtttgtaat gaagacagct cttctctatc tttccataag cctgagatac aagttcaggg      120
actcagcaat gcacttttagg actgagctag gaggcaaata tctgaagctt gctatgctgt      180
tctttccatt ccttttccct ctgaaacaca caaaatacca aaggaactta cgcaacacac      240
cactgagtc tctaactaat catatgtgct cagacacagc tcaagcacac cccttagtta      300
agaaagaacc tccatataca ttaatttttt tctgcctaaa aataaaattg cgttgtggca      360

```


gcaatttgga aactacagca aagtctccaa aaaa 394

<210> 60
 <211> 246
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(246)
 <223> n = A,T,C or G

<400> 60
 cccctccttt tttaggcctg aatacaaagt agaagatcac tttccttcac tgtgctgaga 60
 atttctagat actacagntc ttactcctct cttccctttg ttattcaggg tgaccaggat 120
 ggcgggaggg gatctgtgtc actgtaggta ctgtgccag gaaggctggg tgaagtgacc 180
 atctaaattg caggatggtg aaattatccc catctgtcct aatgggctta cctcctcttt 240
 gccttn 246

<210> 61
 <211> 395
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(395)
 <223> n = A,T,C or G

<400> 61
 ggcacgagct tgcttccctc tcaccctctg cagtttccnn nnnnnnnnnn nnnnnnnnnn 60
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
 nnnnnnnnnn nnnnnncttc catatcgtaa actgccttgg aaccaattac cactaccagg 180
 gagacaaaact attgcttaga ggatgctgac aggagcagca tgccaaaatt ggaagaagga 240
 gaaagttaa gctctcctca ctatgagttt tcaagtataa aagacttttt cttccacgat 300
 tttgagaaca actgaggact cttgtgacca ggacaacagg gaagcttgca gcaagatagc 360
 tccaggttgg attcatgctt cgcaccccaa aggct 395

<210> 62
 <211> 387
 <212> DNA
 <213> Homo sapiens

<400> 62
 ggcacgaggc ttgcggggca ttatgactag agggttggtg aaaattcaga cagaatgtaa 60
 cttgacaaag agaagacagc aacaactgta acaattatct tatgaatatt tgcgaaactc 120
 aaagggatct gattggtgac ctctgggctt tatcaaatca acatcacaac ttctagaaga 180
 aagtcaacct tcatctttta caatagaaat catatgtttt gctaaccat tctatatttag 240
 gctgaaaaca attaagagtt atgggtactt aaaaaaatca ttatgtttat aaaattagtg 300
 atagaaggag catagtgttc tatacagtca cacacataca cttccttatt tcttttatatt 360
 aaactttgag taacatagca gtctatg 387

<210> 63
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 63
 ggcacgaggg aaactgtatg acaggagaat gaatcagggt tgggggtcaa ggtgccggcc 60
 actgggaaaa acagctgccc cgagttgcaa aactctgggt cctatatgta taaactatgc 120
 cctgaggaag gaatctcagg cgtatcttag gagaaaatgt tctagcttgg gaaacaaaca 180
 caacaggacc gtgaatccaa atatttcaag tgggtttaga ggactggagt tctaaacgct 240
 gcttttactg taagtgatca cgccccgaa tgtgctgaag aaaggaaaat gagccagtat 300

cggcgaggac tatgggcaag gaaaacgaga gtgtgcatg tgtcaaagca agacatctgt 360
gtatagtaat ataaccaagt aatagatagt catagaatca a 401

<210> 64
<211> 274
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(274)
<223> n = A,T,C or G

<400> 64
cacgcaccgc cctgtgtgtg tgcgcacaca cgctccctct ctctatagac agacacacac 60
tgcgcgctcg ctctctcttt tgtgtgcgct ctccgtgctc ccccccctctc tctctttttt 120
ctctatatnn nnnnnnnnnn nnnnntctga gagctcgcg cgtcagcgtt ctattcacac 180
gcgcgttttt tttatatata tattttgtgc gcgcggggg gggcgcacac actctctctt 240
ttttgtgggt tcgctgtccg cgtccctctc ttgt 274

<210> 65
<211> 279
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(279)
<223> n = A,T,C or G

<400> 65
cccttttttt tatacacccc cccttgtctg tctgttttgt gtgtctgccc cccttctctc 60
gttgtgatct cctctctctt tttttctccc cccgcgctct ctctctcttg cggggagggt 120
cacatacccc ctctctctct cttttttgaa ccacacattc cgtttctctt ttttttatct 180
ctaccctctc ctctgtctga cccccacan nnnnnnnnnn nnnnnnnnnn nnnagagtag 240
agttgcgttc cccactctcc nnnnnnnnnn gtgggggtgc 279

<210> 66
<211> 311
<212> DNA
<213> Homo sapiens

<400> 66
caaaacaaaa attaaaaatg accccccttt aaaatttttag ggggtccatt tttaaaaacc 60
ttaacagttt aaaggttctt ggtcagtttg gggaacccca ccttgagatg ggagcaaaaa 120
aggggatttt ttccaacat agcgagcggg ttagattttt tttgtccgt tagagttgcc 180
ctgtgcacca cgccaaaacc tccagaggtc ttcttttttt acacaccctg tctgggggtg 240
tttctcagaa gattaacaca gcgcctgggg gtttaaggga ggggtgacct ccgcaggaca 300
ttatggggct t 311

<210> 67
<211> 386
<212> DNA
<213> Homo sapiens

<400> 67
ggcacgaggg aatctcaggc gtatcttatg agaaaatggt ctagcttggg aaacaaacac 60
aacaggaccg tgaatccaaa tatttcaagt gggtttagag gactggagtt ctaaacgctg 120
cttttactgt aagtgatcac gcccgggaat gtgctgaaga aaggaaaatg agccagtatc 180
ggcgaggact atgggcaagg aaaacgagag tgtgcatgt gtcaaagcaa gacatctgtg 240
tatagtaata taatcaagta atagatagtc atagaatcaa gctgatgtat ttggcagggg 300
ccgcgggagg atgaggcaac tcccatcaga ttagaaagat gttaacactg taacaaaagt 360
ggggctcgag gaaggggaaa agcgca 386

<210> 68
 <211> 396
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(396)
 <223> n = A,T,C or G

<400> 68
 ggcacgagga ggcagctgcc tttgtttgcc atggatgggt aggggctgca ctgagcagca 60
 ccggtgttct tcatccggct gcacccccaa cagagctctt tcttccccag atccctttta 120
 cagttggatt ctccctcttg gatctggctc tgccttagtc cgacctagag ggatcagctt 180
 cgcccacgcc cactctcacc cggaaccttt catctcttat tgaagccttt taggcccatt 240
 gggatgttca ttagaactct gaaaactaca gttctccctt ttatgaggac tgcaccacag 300
 ctgcgcctct cctgggttcc gcctgggtgc agagtgcgc catgggacag ccctctgaaa 360
 ttatactgct tacaacatg ctgagtctgc aaggan 396

<210> 69
 <211> 397
 <212> DNA
 <213> Homo sapiens

<400> 69
 ggcacgagtc ttagtcaaca tggacaacaa catcattcag cattacagca accacgtcgc 60
 cttcctgctg gacatggggg agctggacgg caaaattcag atcatcctta aggagctgta 120
 aggctctctg agcatccaaa ccctcacgac ctgcaagggg ccagcagggg cgtggcccca 180
 cgccacacac aacctctcca catgcctcag cgctgttact tgaatgcctt ccctgagggg 240
 agaggccctt gagtccacaga cccacagacg tcagggccag ggagagacct aggggggtccc 300
 ctggcctgga tccccatggt atgcttgaat ctgctccctg aacttcctgc cagtgcctcc 360
 ccgtacccca aaacaatgtc accatgggta ccaccta 397

<210> 70
 <211> 394
 <212> DNA
 <213> Homo sapiens

<400> 70
 ggcacgagcc aaacctagca caaaacgggg ttcacaagcc atgggtcgggg tccggggggg 60
 acagaaatgg attttcttgg caataagcgg actctgggac tccggctccc taccctaaac 120
 tgaagcgctt ccgtgaacac ccccgctctc cgtaggggga ggggagcagg cgggatcctg 180
 ggtccctcat aagcactttg gttttaccgc ctgcaacctc actgtgcccg ccccgcacca 240
 tgccctagcc ccaggtctag ccggggcccat tgcagggggc agcacttggg ggcatctccg 300
 gcacttggtt gggaccaagg agatgccacc atagaccttt ccctcgcctt cttcctccct 360
 agtcggggtt ccattctttt caccagcacc catc 394

<210> 71
 <211> 389
 <212> DNA
 <213> Homo sapiens

<400> 71
 ggcacgagga aagttaagca tctacaggtt atggcttttg gagttccaat atcagtctat 60
 cttttattca acgcaatgac agcactgacc gaagaggcag ccgtgactgt aacacctcca 120
 atcacagccc agcaaggtaa ctggacagtt aacaaaacag aagctgacaa catagaagga 180
 cccatagcct tgaagttctc acacctttgc ctggaagatc ataacagtta ctgcatcaac 240
 ggtgcttgtg cattccacca tgagctagag aaagccatct gcagggtgtct aaaattgaaa 300
 tcgccttaca atgtctgttc tggagaaaga cgaccactgt gaagcctttg tgaagaattt 360
 tcatcaaggc atctgtagag atcagttag 389

<210> 72
 <211> 396

<212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(396)
 <223> n = A,T,C or G

<400> 72
 ggcacgaggc ctggcccccac agcggggcag caggatctct gtgctgtcag ccagcccagt 60
 gtctgatgtc agctatatgt ttggaagcag ccagtcacct ctgcactcca gcaactccag 120
 ccatcagtca tcttccagat ccttggaaag tccagccaac tcttcctcca gcctccacag 180
 ccttggctca gtgtccctgt gtacaagacc cagtgaattc caggctccca gaaaccccac 240
 cctaaccatg ggccaaccca gaacaccca ctctccacca ctggccaaag aacatgccag 300
 cagctgcccc ccatccatca ccaactccat ggtggacata cccattgtgc tgatcaacgg 360
 ctgcccagaa ccagggtctt ctccacccca gcggan 396

<210> 73
 <211> 386
 <212> DNA
 <213> Homo sapiens

<400> 73
 ggcacgaggc cacctgttgc cctaacaccc tgtctgactc tctcccgtcg cagcagccag 60
 tccctcctgc actccagcaa ctccagccat cagtcactct ccagatcctt ggaaagtcca 120
 gccaaactct cctccagcct ccacagcctt ggctcagtgt ccctgtgtac aagacccagt 180
 gacttccagg ctcccagaaa ccccacccta accatgggcc aaccagaaac accccactct 240
 ccaccactgg ccaaagaaca tgccagcagc tgcccccat ccatacacia ctccatgggtg 300
 gacataccca ttgtgtgat caacgggtgc ccagaaccag ggtcttctcc accccagcgg 360
 accccaggac accagaactc cgttca 386

<210> 74
 <211> 390
 <212> DNA
 <213> Homo sapiens

<400> 74
 ggcacgagct cagatccggg gactgcggat aaatggcctt aggcgcgggg cagcgagatg 60
 ttgcgttccg gtgtgggtgt ggggtgtcct ccgacggcgt ctcggtgcca gtgtcgaggt 120
 tctttctgct tagctacccg gagccgacta cggaggagga cacctgagtt tacgtctctt 180
 ccatctgctg ctgcctcag ctgcctgggt ccccgacgag agccaggtga cacttaactc 240
 cgccatctgc gttttgagca ctgttctcat aatggagttt cctgatttgg ggaagcattg 300
 ttcagaaaag acttgcaagc agctagattt tcttccagta aaatgtgatg catgtaaaca 360
 agattttctgt aaagatcatt ttccatacgg 390

<210> 75
 <211> 399
 <212> DNA
 <213> Homo sapiens

<400> 75
 ggcacgagaa atggccttag gccgcgggca gcgagatggt gcgttccggt gtgggtgtgg 60
 gtgtgacctc gacggcgtct cggtgccagt gtcgaggttc tttctgctta gctacccgga 120
 gccgactacg gaggaggaca cctgagttta cgtctcttcc atctgctgct cgctcagct 180
 gcctgggtcc ccgacgagag ccaggtgaca cttaactccg ccattctgctg tttgagcact 240
 gttctcataa tggagtttcc tgatttgggg aagcattggt cagaaaagac ttgcaagcag 300
 ctagattttt ttccagtaaa atgtgatgca tgtaaacaa atttctgtaa agatcatttt 360
 ccatacgctg cacataagtg tccgtttgca ttccagaag 399

<210> 76
 <211> 386
 <212> DNA
 <213> Homo sapiens

<400> 76
 ggcacgagca aaggctcgca ggggccagaa acccggtccc gagcggcggc ggcccggctt 60
 ccgctgcccc tgagctaagg acggtccgct ccctctatcc agctccgaat cctgatccag 120
 gcggggggcca gggggcccctc gcctcccctc tgaggaccga agatgagctt cctcttcagc 180
 agccgctctt ctaaaacatt cataccaaag aagaatatcc ctgatggatc tcatcagtat 240
 gaactcttaa aacatgcaga agcaactcta ggaagaggga atctgagaca agctgctatg 300
 ttgcctgagg gagaggatct caatgaatgg agtgctgcga acacctgggg attcttttac 360
 cagcaacaac atggtttttg ggaact 386

<210> 77
 <211> 395
 <212> DNA
 <213> Homo sapiens

<400> 77
 ggcacgaggc catctccaaa tactgcggtt gttcagaagc tcttagtttg tgggctgtcc 60
 ttgttatctt acttgaccat ctgtacaaca ttacctgtgg agtacaacat tgatgagcat 120
 tttcaagcta cagcttcgtg gccaaacaaag attatctatc tgtatatctc tcttttggtc 180
 gccagaccga aatactatct tgcattggacg ctactgtatg ccattaataa tgctgcaggc 240
 tttggtttca gaggggtatga cgaaaatgga gcagctcgct gggacttaat ttccaatttg 300
 agaattcaac aaatagagat gtcaacaagt ttcaagatgt ttcttgataa ttggaatatt 360
 cagacagctc tttggctcaa aaggggtgtg tatga 395

<210> 78
 <211> 389
 <212> DNA
 <213> Homo sapiens

<400> 78
 ggcacgaggc aggcgggat gttcgtcctg gtggaaatgg tggacaccgt ccgatcccc 60
 ccttggcagt ttgagaggaa gctcaacgac tccattgccg aggagctgaa caagaagttg 120
 gccaaacagg tcgtgtacaa cgtgggactc tgcatttgctc tgtttgatat caccaaactg 180
 gaggatgcct atgtattccc tggggatggc gcatcacaca ccaaagtcca ttttcgctgc 240
 gtggtgtttc atccattcct agatgagatt ctcatggga agatcaaagg ctgcagccca 300
 gaaggagtgc acgtctctct aggcctcttc gatgacattc tcatccccc agagtcactg 360
 cagcagccag ccaagttcga cgaagcgga 389

<210> 79
 <211> 365
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(365)
 <223> n = A,T,C or G

<400> 79
 ggcacgagaa aacatttcat cttgattttt attaaggatga tatgtatggt acttaacagc 60
 tgtataatac acatttgcatt gcattaggaa gttttttttg ggtttttattc atcctgtagt 120
 gatgtatctg tgacctcaac gagtaggcac ttctgtactg tactggtttc ttaaagtttc 180
 ttttatcccg ccccccacccc caacctcagc ctcaagtatg taannnnnnn nnnnnnnnnn 240
 nnnnnnnnnn nnnnnnnnnn nnnnnnaaac aaagccccg tttgtcccca ggctggataa 300
 caggggcgga atctgggtta attgaaccct ttgctttttg gggttaaggca attttcctgc 360
 ctcac 365

<210> 80
 <211> 376
 <212> DNA
 <213> Homo sapiens

<400> 80
 ggcacgagct ggaaaccagc ccctaagctg ctttcacctt cggcccattt ctcaactagcc 60

agcctctccc	acggcctccc	cagtttcttc	aaatacaccc	cctccactat	tcaccatact	120
gccaccgtga	tttatattaca	actttttgtc	cggattacct	cagtagcctt	ctaattgtcc	180
ccttttgcac	taaagtagcc	cctctcatcc	cccaaactct	accatgtcac	tcttctacat	240
aattctggct	ttccatgacc	cataaaccac	atctctcaag	tgtgctctat	gctggcttga	300
atatgttaat	gatcttaatt	ctacttttag	tgcaattttc	ttagagctgg	catcactttc	360
atcatgacgt	gagaac					376

<210> 81

<211> 384

<212> DNA

<213> Homo sapiens

<400> 81

ggcagcagag	gattgtgtga	aattgtgcaa	atgcatgaat	gtgggctggg	atagtaaaag	60
ggaggggccc	ggagcagccc	acctggggtc	ctatctagta	gacgcgcccg	gtgcccaccc	120
attgctgtga	tgccagcagc	ccactgcaag	catcctcttc	ctttccaagg	ttctgtctgg	180
tacatgaata	ggtgtggcag	gggtgggggc	tcctgaagac	caactagggg	tactagggac	240
cttagactct	tgcgagagcc	tgcaccccat	atcaggtggg	gtcaatagat	aaataccccc	300
gcctccttgc	cccttagttc	tgggtgtggtg	ggcaagtcag	aggaactggt	cttctcacac	360
tttcacgtgc	tctcgttga	gatc				384

<210> 82

<211> 383

<212> DNA

<213> Homo sapiens

<400> 82

ggcagcagca	aaggctcgca	gcggccagaa	acccggctcc	gagcggcggc	ggcccggcctt	60
ccgctgcccg	tgagctaagg	acggtccgct	ccctctagcc	agctccgaat	cctgatccag	120
gcggggggcca	ggggcccctc	gcctcccctc	tgaggaccga	agatgagctt	cctcttcagc	180
agccgctctt	ctaaaacatt	caaaccaaaag	aagaatatcc	ctgaaggatc	tcatcagtat	240
gaactcttaa	aacatgcaga	agcaactcta	ggaagtggga	atctgagaca	agctgttatg	300
ttgcctgagg	gagaggatct	caatgaatgg	attgctgtga	acaactgggg	atttctttac	360
caggatcaca	atggtaatat	ggg				383

<210> 83

<211> 358

<212> DNA

<213> Homo sapiens

<400> 83

ggcagcagca	gggccgcgcg	gcggtgatca	agcacccgctt	ccccaagggc	taccggcacc	60
cggcgctgga	ggcgcggtt	ggcagacggc	ggacgggtgca	ggaggcccgg	gcgctcctcc	120
gctgtcgccg	cgctggaata	tctgccccag	ttgtcttttt	tgtggactat	gcttccaact	180
gcttatatat	ggaagaaatt	gaaggctcag	tgactgttcg	agattatatt	cagtccacta	240
tggagactga	aaaaaactcc	ccagggtctc	tccaacttag	ccaagacaat	tgggcagggt	300
ttggctcgaa	tgcacgatga	agacctcatt	catgggtgac	tcaccacctc	caacatgc	358

<210> 84

<211> 338

<212> DNA

<213> Homo sapiens

<400> 84

aagatggctg	agagggacag	aatgctttat	tttggagaga	aacaatgttc	taggtcaaac	60
tgagtctacc	aaatgcacac	tttcacaatg	ggtctagaag	aaatctggac	aagtcttttc	120
atgtggtttt	tctacgcatt	gattacatgt	ttgctcacag	atgaagtggc	cattctgcct	180
gcccctcaga	acctctctgt	actctcaacc	aacatgaagc	atctcttgat	gtggagccca	240
gtgatcgcg	ctggagagac	agtgtactat	tctgtcgaat	accaggggga	gtacgagagc	300
ctgtacacga	gccacatctg	gattcccagc	agctgggtg			338

<210> 85

<211> 475

<212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(475)
 <223> n = A,T,C or G

<400> 85
 gtcgctcaat aggcaggagt ccatcgattc gaattcggca cgagnnnnnn nnnnnnnnnn 60
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 120
 nnnnnnnnnn gctccactgt gcactcctga cacatacttt ccccgctaca ctctctattc 180
 tccccctctt gtgttctctc tctatagcgg tagatagaga ggcctgtgtg tagataataa 240
 acgtgtgtgt gtgtgtaaga aaggagacac aaacacgccc acnnnnnnnn nnnttggggc 300
 ctttttttct tttgagccct ttggggaaaa aaccgggga aaacagccca taccactat 360
 ttggggcgcg ccaaaaaaacc ttctttaaaa aaaatgtgtt aaatgttaaa ttttttagga 420
 acannnnnnn nnnngcaaaa aatagcacc caaaagcagg ggttttacat ttttg 475

<210> 86
 <211> 467
 <212> DNA
 <213> Homo sapiens

<400> 86
 gagcgatttt ctgcaggatt ctatcgattc gaattcggca cgagccatgg tctcagttag 60
 ggctggaatt tacagagaag tttggccagg ggtgccacca tgctgccagt cagtttggga 120
 aggaacaga gaagctcggc catgggggtcc accatggggt taatgaggcc tggaaggaag 180
 cagagaagtt tggccagggt gtccaccatg ctgcctcgca ggtggggaag gaggaagaca 240
 gagtgggtcca aggcctccat catggcggtta gtcaggctgg aaggaggagg gggcagtttg 300
 gccacgacat tcaccacaca gcagggcagg ctgggaaaga gggagacata gcagttcatg 360
 gtgtccaacc tgggggtccac gaggcgggga aggaggcagg gcaatttggc cagggagttc 420
 accataccct tgaacaggcc gggaagggaag caaacaagc ggtccag 467

<210> 87
 <211> 449
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(449)
 <223> n = A,T,C or G

<400> 87
 cggggtggga aaccngannt tnannaancg gacggattct cccgttccga atagcctttt 60
 acagaagatt cttcacagct atgtgcctga agagatcang gatggaaatc aagttcgagt 120
 tacctcatgg gatggcagga aatggggaga actggagggg gacacctatg accgggtgct 180
 ggtgatgtg ccctgtacca cagaccgcca ctcccttcat gaggaggaga acaacatctt 240
 taagcgggtca aggaagaagg agcgacagat attgcctgtg ctgcaagtgc agcttcttgc 300
 ggctggactc cttgccacca aaccaggagg ccatgttgtc tattctacct gctcactctc 360
 acacttacag aacgagtatg tgggtcaagg tgccattgag ctccctgggca atcaatacag 420
 catccaggta caggtggaag atctgactg 449

<210> 88
 <211> 439
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(439)
 <223> n = A,T,C or G

```

<400> 88
gtagtgtatg tgcagcctcc catcgattcg aattcggcac gagatccccct cttatatgat      60
gccccagccc aggagagata aaagcatcag caccatgaga ttcacctgcc tctgggtcgtn      120
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      180
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      240
nnnnnnnnnn actcttagac agcaaaaatg ctttctccca gtcttgttcc cttgtttctca      300
gttcccacc tgcctggata actactgttc ttggttttnn nnnnnnnnnn nnnnnnnnnn      360
nagtctcgta ccagattcaa aaatcagtc actacttcaa aaacaatgac atgctggcta      420
cttagataga agaggaggc                                     439

```

```

<210> 89
<211> 436
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(436)
<223> n = A,T,C or G

```

```

<400> 89
ggcagcagca tcaaatagta aatatagatc ttatgctgga aatgtcaacc tccctggcag      60
ctgtaacgcc catcattgaa agggaaagcg gaggacacca ttatgttaat atgactttac      120
ctgtcgatgc agttatatct gttgctccag aagaaacatg gggaaaagtt cgtaaactcc      180
tggttgatgc aattcataat caactaactg acatggaaaa atgtattttg aaatatatga      240
aaggaacatc tattgtggtc cctgaaccac tgcacttttt attaccaggg aaaaaaatc      300
ttgtaacaat ttcatatcct tcaggaatac cagatggcca gctgcaggcc tataggaagg      360
agttacatga tcttttcaat ctgcctcacg acagacccta tttcaaaagg tctaattgctt      420
atcacttttc agatgn                                     436

```

```

<210> 90
<211> 437
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(437)
<223> n = A,T,C or G

```

```

<400> 90
ggcagcagag atcatgcact accacatgca gcacgagcag taccggcagg tcatcagcgt      60
gtgtgagcgc catggggagc aggaccctc cttgtgggag caggccctca gctacttcgc      120
tcgcaaggag gaggactgca aggagtatgt ggcagctgtc ctcaagcata tcgagaacaa      180
gaacctcatg ccacctcttc tagtggtgca gacctgggcc cacaactcca cagccacact      240
ctccgtcatc agggactacc tggtcacaaa actacagaaa cagagccagc agattgcaca      300
ggatgagctg cgggtgcggc ggtaccgaga ggagaccacc cgtatccgcc aggagatcca      360
agagctcaag gccagtccta agattttcca aaagaccaag tgcagcatct gtaacagtgc      420
cttggagttg ccctcan                                     437

```

```

<210> 91
<211> 437
<212> DNA
<213> Homo sapiens

```

```

<400> 91
ggcagcagct tcagtcttat gtcatttact ctttaggaca acctcttgaa aaactaaatc      60
atctctttga aggtgttgaa gctcgcgtgg cacagggcat aaggaggag gaagtaagtt      120
accaacttgc atttaacaaa caagaacttc gtaaagtcac taaggagtac cctggaaaag      180
aagtaaaaaa aggtctagat aacctctaca agaaagtga taaacattta tgtgaagaag      240
agaacttact tcagggtggtg tggcactcca tgcaagatga atttatacgc cagtataagc      300
actttgaagg tttgatagct cgctgttatc ctggatctgg tgttacaatg gaattcacta      360
ttcaggacat tctggattat tgttccagca ttgcacagtc ccactaaacc ttgtgaaaga      420

```


agaaaagata actgaat

437

<210> 92

<211> 427

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(427)

<223> n = A,T,C or G

<400> 92

aacggctctt	ctncttttga	ggagcccatc	gagtcgaatt	cggcacgagg	cgagtctctg	60
ggtcgcgacg	ggaaggagtg	aaacacctct	ctgcgcctgc	gcgctccgtg	cctgcgaagc	120
aaacccggcc	tcaccttttc	ctgcccgaag	cagaagattc	tcgcaggcct	ggtttctccc	180
tccagaagac	ccccaccca	aatcctctgt	agctcctggg	agtgccctga	cccctgctgc	240
caccgtcctt	cagagagcaa	cgaagagct	tcccggaggg	cgaggaaaag	agggaaagta	300
gccagcaatg	tcgaacgcag	tgtataataa	gatgtggcat	cagacccaag	aagccctcgg	360
tgctttactc	gatgaagagc	ctcagacgat	gattgaacca	cacagaaatc	aggttttcat	420
ctttcaa						427

<210> 93

<211> 429

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(429)

<223> n = A,T,C or G

<400> 93

gtgacgatcc	catcattcaa	ttcggcacga	gctcacagcc	aaagtccctt	ctgccccag	60
gctgagagtg	cttgatacac	ccttgaatcc	cctcttatat	gatgccccag	cccaggagag	120
ataaaagcat	cagcaccatg	agattcacct	gcctctggtc	gttaggggaa	aatggaggcc	180
tgcgatttgg	agttaaactc	tcagtgatct	ctgtgttgac	aacaccaaag	ctagaggaat	240
ccagtaggat	gtgggcatgg	ttttcccgga	aggctgactg	agcagttctg	caaagtgttg	300
caagtacagg	gcagaatttc	atccagcctc	agaaccttga	gccaagactc	agcatcagca	360
aagccaaaag	tttcatttct	tcgactgtgg	gagtgtctagt	cccaaccttt	agatggccat	420
tcagttnta						429

<210> 94

<211> 421

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(421)

<223> n = A,T,C or G

<400> 94

ggcacgagat	tattttacttg	gtgtgtggtc	accactgttt	tttaaagtga	tgttttcatt	60
tgtatcaaac	tggacctgct	ttcctcaagg	attgcccaca	aggagacaca	aattttactaa	120
acacttatca	ataatagaac	accgtgctag	gcaatttcca	tatactatta	atttaatcct	180
cacaataact	ttggaagaca	gaaagtattt	tctctgannn	nnnnnnnnnn	nnnnnnnnnn	240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	300
nnnnnnnnnn	atcctctgtc	tccaaagcct	gtacttcatt	caggacactt	tccccacat	360
ttagaaaagc	tgtaattatc	ttccagtgag	acagcatagc	acatgtgatc	actgtccctt	420
c						421

<210> 95

<211> 421
 <212> DNA
 <213> Homo sapiens

<400> 95
 ggcacgagat gagaagataa aattcagcgt tggccttttag actttgccat ccttaaggag 60
 tgatggaagc caagtgaaca agcctcagtg acacaagtca aattcatagt ttcactcttg 120
 gttttttgtt gttgtgtgtt tattattctc actacagaaa gactgagttt catgctcctg 180
 gctatgtcag atgtgaattt tcatgggtaa ctggacagtt aacaaaacag aagctgacaa 240
 catagaagga cccatagcct tgaagttctc acaccttttg ctggaagatc ataacagtta 300
 ctgcatcaac ggtgcttgtg cattccacca tgagctagag aaagccatct gcagggtgtt 360
 tactggttat actggagaaa ggtgtctaaa attgaaatcg ccttacaatg tctgttcttg 420
 a 421

<210> 96
 <211> 418
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(418)
 <223> n = A,T,C or G

<400> 96
 tggatccatc gattcaattc ggcacgaggt .tatttttaag aacttttgct tactatatattg 60
 gattttacctg cgggtgtgagt agcttttaaat gtttgtgttt atacagataa gaaatgctat 120
 ttctttcttg ttctgtcagc cattgaaaaa cctttttcct tgcaaattat aatgtttttg 180
 atagattttt atcaactgtg ggaaaccaa cacaagctg ataacctttc ttaaaaacga 240
 cccagtcaca gtaaagaaga cacaagannn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 360
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnng 418

<210> 97
 <211> 418
 <212> DNA
 <213> Homo sapiens

<400> 97
 atgctacatt gctactttgt tgcattgatc gccagacgac cactagattc gaaactgagc 60
 gcgagatgat gaatctgtgt gttatgaaaa tgcattgtac cagtataaaa tgacattctc 120
 tattaataac atctgcggtg cgacacacat aattgtccca attttttaata ttgatgggga 180
 gcatgaagca tttttttaat gtgttggcag gccccattaa atgcataaac tgcataaggac 240
 tcatgtgggtc tgaatgtatt ttagggcttt ctgggaattg tcttgacaga gaacctcagc 300
 tggacaaagc agccttgatc tgagtgtgct aactgacaca atgaaactgt caggcatgtt 360
 tctgtctctc tctctggctc ttttctgctt ttttaacaggt gtcttcagtc agggaggg 418

<210> 98
 <211> 417
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(417)
 <223> n = A,T,C or G

<400> 98
 catcgattcg aattcggcac gaggccaagt ggacaggcca tagcccccac agactggagg 60
 gacgcggcta gggaatgtcc cacagagtgg ccagttatcc ctgagagaaa gagcagggtt 120
 tagcggagac tctgaggctg ctttagaata tgggtgggtgt gtggggcaaa agggacaccc 180
 aggggtgtat caagaggtca tnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 240
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 300

nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnn 360
 nnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnn 417

<210> 99
 <211> 416
 <212> DNA
 <213> Homo sapiens

<400> 99
 ggcacgagct acctcagccc tgctccagga gacaccagca gctggggccag tggccctgag 60
 agatggcccc gaagggagca tgtggtgaca gtcagcaaga ggaggaacac atctgtggac 120
 gagaactatg agtgggactc agaattccct ggggacatgg aattgctgga gactttgcac 180
 ctgggcttgg ccagctcccg gctcagacct gaagctgagc cagagctagg tgtgaagact 240
 ccagaggagg gctgcctcct gaacactgcc catgttactg gccctgaggc ccgctgtgct 300
 gcccttcggg aggaattcct ggcttccgc cgccgccgag atgctactag ggctcggcta 360
 ccagcctatc gacagccagt cccccacccc gaacaggcca ctctgctgtg aacatt 416

<210> 100
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 100
 ggcacgaggg aaaatgtagg ctaccagtag aaaatgacat tctctattaa taagatctga 60
 ggtgcgacac acataattgt cccaattttt aagattgatg gggagcatga agcatttttt 120
 taatgtgttg gcaggcccca ttaaatgcat aaactgcata ggactcatgt ggtctgaatg 180
 tatttttaggg ctttctggga attgtcttga cagagaacct cagctggaca aagcagcctt 240
 gatctgagtg agctaactga cacaatgaaa ctgtcaggca tgtttctgct cctctctctg 300
 gctcttttct gctttttaac aggtgtcttc agtcaaggag gacagggttg ctgtggtgag 360
 ttccaggaca ccaaggtcta ctgcactcgg gaatctaacc cacactgtgg ctctgat 417

<210> 101
 <211> 412
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(412)
 <223> n = A,T,C or G

<400> 101
 ggcacgagga aagtaaactgt gtatctcttg ttcattttta tagaactttt gcatactata 60
 ttggattttac ctgcgggtgtg actagcttta aatgtttgtg tttatacaga taagaaatgc 120
 tattttctttc tgggttcctgc agccattgaa aaaccttttt ccttgcaaat tataatgttt 180
 ttgatagatt tttatcaact gtgggaaacc aaacacaaag ctgataacct ttcttaaaaa 240
 cgaccagtc acagtaaaga agacacaaga nnnnnnnnnnn nnnnnnnnnnn 300
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn 360
 nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nnnnnnnnnnn nn 412

<210> 102
 <211> 414
 <212> DNA
 <213> Homo sapiens

<400> 102
 ggcacgaggt cttgctcaca tgttgacta ctctctcctg gatgtcactt gtcacctcta 60
 ccagccctcc tttctccaga tggtctcttc ataaccacca ggtcagaaga ggatccgttc 120
 caatgatttt cctaaaacaa tggaagtgtt ttccaaagag cttataaggc attgtaggat 180
 ctggcctgcc ctgactccac tttaccagaa ccatctgctg ctcttctctc ttgtgttact 240
 caaggtatta gctgctgtgg caaatcaact ctgaaatctc cgtgacttaa tacaagagag 300
 gtttattttct tactcacgct ggggtgactg ccacttggtg acagaggagc tatggaaact 360
 tgagacctaa gcagaaatga gttcaataat attgctacac tctaggactt tctc 414

<210> 103
 <211> 410
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(410)
 <223> n = A,T,C or G

<400> 103
 ggcacgagga agagccggga ggatgtattg gttgttagga aaatgtaggc taccagtaga 60
 aaatgacatt ctctattaat aagatctgag gtgcgacaca cataattgtc ccaattttta 120
 agattgatgg ggagcatgaa gcattttttt aatgtgttgg caggcccat taaatgcata 180
 aactgcatag gactcatgtg gtctgaatgt atttttagggc tttctgggaa ttgtcttgac 240
 agagaacctc agctggacaa agcagccttg atctgagtga gctaactgac acaatgaaac 300
 tgtcaggcat gtttctgctc ctctctctgg ctcttttctg ctttttaaca ggtgtcttca 360
 gtcaaggagg acaggttgac tgtggtgagt tccaagacac ccaaggctan 410

<210> 104
 <211> 411
 <212> DNA
 <213> Homo sapiens

<400> 104
 ggcacgagat acgaatgggg tgtatttttc gactgctcgc aggcaccccc aggttatgtg 60
 gacagagcta agcccaaagt tgtgattttc cactctgttc tgtccatgtc gagggagat 120
 aagtagaaag tgacacagta agagccagaa tacaccaggt gaaggagaga attgcattgt 180
 gttttgagaa gtttctactga caagtatatcc tgggctgtgg gacatcacta gctttgaaag 240
 tgtagctggc acctcgctca tctaatttga tgggtgtgtg tggggtgttg tgcacgcgtc 300
 ggtctaacat atctgaaccc aggtgatttc tgttctcagg acgcttttag gtgacaagga 360
 tcaggcatgt gaacaaataa ccatactgta aagctggctg tgctgggtct c 411

<210> 105
 <211> 413
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(413)
 <223> n = A,T,C or G

<400> 105
 ggcacgagga agattctcgc agtcctggtt tctccctcca gaagaccccc caccctaatg 60
 ctctgtagct cctggtagtg ccctgacccc tgctgccacc gtccttcaga gagcaacgga 120
 agagcttccc ggagggcgag gaaaagaggg aaagtagcca gcaatgtcga acgcaatgta 180
 taataagatg tggcatcaga cccaagaagc cctcggtgct ttactcgata aagagcctca 240
 gaagatgatt gaaccacaaa gaaatcaggt tttcatcttt caaacattag ccaccttcta 300
 cgtaaagtat gtgcagatct ttagaaacct agagaatgtc tacgaccagt tcgtccaccc 360
 ccagaaacga atactgatca ggaaagtctt ggacgngtg atgggccgca tcc 413

<210> 106
 <211> 412
 <212> DNA
 <213> Homo sapiens

<400> 106
 aggatcccat cgatttcta tctggcacgag ctccataagg cagaggtcta tgcgaggacg 60
 cccggctgga ccacgagacc gccattgat tgcgctggga caagaattcc ttatctttgg 120
 aggcagtga acgactaata gctaaaggta atacagaaga actacgaaaa tgttttgggg 180
 tccgaatgga gtttgtgaca gctggcctcc gagctgctat gggacctgga atttctcgta 240
 tgaatgactt gaccatcatc cagactacac agggattttg cagatacctg gaaaaacaat 300

tcagtgaactt	atagcagaaa	ggcatccgga	tcagttatga	cgcccagagct	catccatcca	360
gagggggtag	catcaaaaagg	tttgcccagac	ttgctgcaac	cacattttatc	ag	412

<210> 107
 <211> 408
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(408)
 <223> n = A,T,C or G

<400> 107						
ggcacgagga	aaaaccagtt	tctctttttat	tgtctgttac	taatctctat	tctaaagatt	60
cagctcaatt	ctcaaccata	ctccaaactc	tctctttttcc	agctaccttt	actccctctc	120
cttcaattcc	actttcctct	gcttacnnnn	nnnnnnncnn	nnnnnnnnnn	nnnnnnnnnn	180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnngggn	240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	naatgttttt	360
tttcatttaa	gagagaaatc	acctattcag	gaccggcccc	cacctttg		408

<210> 108
 <211> 405
 <212> DNA
 <213> Homo sapiens

<400> 108						
ggcacgaggc	ttacaggggt	gaccagggcc	cttcctaact	cgaccgcatg	tggattgggtg	60
gctggcttgg	gagggagggt	gtccgatgct	gacattcccc	ttaacatggc	cctgaccgtg	120
gctgtcaggg	gccaccttgc	ctcaccaggc	cagccccact	gggaatgggg	tcagtcacag	180
cagaaccgtc	caaaggtgga	cctgatgtgg	gccctgcccg	gggcgcttgg	cctcagcggg	240
ccatgggaga	cccagtgaac	cgactctagt	gtgaggcagt	ggtcctgcca	ctgactgaca	300
aaccctcttt	gtaagcaaac	ttgacaaata	atgaatctac	tgaactctgt	tatagaacaa	360
gctcattctg	catgaacttc	tcttattgaa	gcagaagcca	cgtca		405

<210> 109
 <211> 403
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(403)
 <223> n = A,T,C or G

<400> 109						
ggatcccata	gnttcgnatt	cggcacgagg	caaccagctc	gtccagcgcg	tggccctgct	60
gctcaaggag	cagactgcgt	acccccgac	acactacatc	cggaggggtgc	cccagaggaa	120
gatccactac	ttcacgggcc	tgcaggcgct	tcagctgctg	ctgctgtgtg	ccttcggcat	180
gagctccctg	ccctacatga	agatgatctt	tcccctcact	atgatcgcca	tgatcccat	240
ccgctatatc	ctgctgcccc	gaatcattga	agccaagtac	ttggatgtca	tggaactgta	300
gcacaggcct	tgactggcag	accctgcccc	cgccccattc	gccagccctc	cacgtactcc	360
caagctggct	ctggaactgt	gaggggaagg	ggaagatgtg	tgg		403

<210> 110
 <211> 397
 <212> DNA
 <213> Homo sapiens

<400> 110						
ggcacgagtc	tgcttctgtc	actgtcacat	agacagccct	gcattgcccc	tgtctcacac	60
aggttgtaat	gaagacagct	ccttctcact	tttcataaag	cctgagatac	aagttcaggg	120

actcagcaat	gcacttttagg	actgagctag	gaggcaaata	tctgaagctt	gctatgctgt	180
tctttccatt	ccttttccct	ctgaaacaca	caaaatacca	aaggaactta	cgcaacacac	240
cactgagtc	tctaactaat	catatgtgct	cagacacagc	tcaagcacac	cccttagtta	300
agaaagaacc	tccatataca	ttaatttttt	tctgcctaaa	aataaaaattg	cgttgtggca	360
gcaatttgga	aactacagca	aagtctccaa	aaaaatc			397

<210> 111
 <211> 401
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 111						
ggcagcagag	ccgttgcctt	caccgccctt	tctcctttta	tcctttttta	aacgctcttg	60
ggggttatgt	ccgtgcttc	ttgggtgccg	agacatatag	atgggtggtct	cgggccagcc	120
cctcctctcc	ccgccttctg	ggaggaggag	gtcacacgct	gatgggcact	ggagaggcca	180
gaagagactc	acaggagcgg	gctgccttcc	gcctggggct	ccctgtgacc	tctcagtcct	240
ctggcccggc	cagccaccgt	ccccagcacc	caagcatgca	attgcctgtc	ccccccggcc	300
agcctcccca	acttgatgtt	tgcgttttgt	ttggggggat	atttttcata	attattttnn	360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	c		401

<210> 112
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 112						
ggcagcaggg	cagtccagca	acaagccttt	catttacatt	aaattataac	ttttcattca	60
ttcctaaacc	aaacttaaaa	ttctgctttc	ctttgagtag	aaggatatta	acttgttttg	120
tttttccttc	agaaggaatt	taatgcaaac	ggattgcagt	cagcactttc	tgaatgtttt	180
cacacagtat	gcaaagctta	catcatacca	aggagtggag	agttgaagtt	tcctcccagt	240
gactccagtg	acagaccaca	cctagaaagc	gtttctcttc	ctgagtattt	caaaaagatg	300
taaaagagct	ggggagagta	tgggaagaaa	caatacagga	ttgcctttaa	ttaattaaga	360
attgcctcct	gataaaagga	aaaagaaatt	aatgctggag	g		401

<210> 113
 <211> 401
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 113						
ggcagcaggg	cccacggggc	ccatctcccc	acaggcattg	agggtaactg	gggtaggctc	60
ctggagcagg	tgggcaccat	ggctttgtgg	gccagccaaa	gggaaaagga	ggtgcttagg	120
agggaaaggg	cagtggaaatg	gcgggagagg	gctgtggaaa	aaagggagcg	agccctggag	180
gaggtggaaa	gggccatcct	ggagatgaag	tgggaaggtga	gggctgagaa	ggaggcatgc	240
cagcggggaga	aagagctgcc	tgcagcagta	catcccttcc	attttgttta	aattgggctt	300
ggagaatcta	ttctgaaaac	attgactcta	gacttgtaga	anagagccat	tttaattttc	360
accttcaatg	gtaaaaagcaa	gggtaatttg	gttgacattt	t		401

<210> 114
 <211> 399
 <212> DNA
 <213> Homo sapiens

```

<400> 114
ggcacgagag  cagaagattc  tctcagtcct  ggtttctccc  tccagaagac  cccccaccca      60
aatcctctgt  agctcctggt  agtgccctga  cccctgctgc  caccgtcctt  cagagagcaa     120
cggaagagct  tcccggaggg  cgaggaaaag  agggaaagta  gccagcaatg  tcgaacgcaa     180
tgtataataa  gatgtggcat  cagaccaag  aagccctcgg  tgctttactc  gataaagagc     240
ctcagaagat  gattgaacca  caaagaaatc  aggttttcat  ctttcaaaca  ttagccacct     300
tctacgtaaa  gtatgtgcag  atctttagaa  acctagagaa  tggctacgac  caggtcgtcc     360
acccccagaa  acgaatactg  atcaggaaag  tcctggacg

```

```

<210> 115
<211> 399
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(399)
<223> n = A,T,C or G

```

```

<400> 115
ggcacgaggc  tttttccaac  ttttaaggat  atcaggagag  aagacactct  tgatgtggag      60
gtttctgcca  gtggctacac  aaaggaaatg  caggcagatg  atgaactgct  tcatccatta     120
ggtccagatg  ataaaaatat  tgaaacaaaa  gagggatctg  aattctcatt  ttcagatgga     180
gaagtggcag  aaaaagcaga  ggtttacagg  tcagaaaatg  aaagtgaacg  gaactgtcta     240
gaagaatcag  agggctgcta  ttgcagatca  tctggagacc  ctgaacaaat  aaaggaagac     300
agtttatcag  aagagagtg  tgatgcacgg  agttttgaaa  tgactgaatt  caatcaagct     360
ttataagaaa  taaaagggca  ggttgttgaa  aacaactcn

```

```

<210> 116
<211> 400
<212> DNA
<213> Homo sapiens

```

```

<400> 116
ggcacgagcg  gaccggggccg  agccggggccg  cccggggcgca  gtctttaacc  atggcggtccc      60
tcttcaagaa  gaaaactgtg  gatgatgtaa  taaaggaaca  gaatcgagag  ttacgaggta     120
cacagagggc  tataatcaga  gatcgagcag  ctttagagaa  acaagaaaaa  cagctggaat     180
tagaaattaa  gaaaatggcc  aagattggta  ataaggaagc  ttgcaaagtt  ttagccaaac     240
aacttgtgca  tctacggaaa  cagaagacga  gaacttttgc  tgtaagtcca  aaagttactt     300
ctatgtctac  acaaacaaaa  gtgatgaatt  cccaaatgaa  gatggctgga  gcaatgtcta     360
ccacagcaaa  aacaatgcag  gcagttaaca  agaagatggg

```

```

<210> 117
<211> 402
<212> DNA
<213> Homo sapiens

```

```

<400> 117
ggcacgaggg  gagatcgctc  agctggccgt  gtcctggcag  gccacggcat  atgcctccaa      60
ggacgggggtc  ctcaactgag  ccatgatgga  cgcctgtgtg  caagatgctg  tccagcagta     120
ccgacagaag  atgcgctggc  tgaaggcgga  ggggcctggg  cgcggggtcg  agcaccacct     180
atccggagtc  caaggcgaga  ccctcacctc  atggagcctg  gccacggacc  cctcctaccc     240
ctgccttgcc  ggccccctgca  catttaggat  atgctcctgg  atggggactg  ggctgtgccc     300
agggcctctg  tccccagga  tgtcttgtgg  tggcggtcgg  ccgttctgcc  cccaggggca     360
ccccctgttg  taggcactgg  ctctaggagg  gcaggcctcc  tt

```

```

<210> 118
<211> 395
<212> DNA
<213> Homo sapiens

```

```

<400> 118
ggcacgaggt  agagatacga  atgggggtgta  gtagccgact  gctcgcaggc  acccccaggt      60

```

tatgtggaca	gagctaagcc	caaagttgtg	attttccact	ctgttctgtc	catgtcgagg	120
gaagataaagt	agaaagtgac	acagtaagag	ccagaataca	ccaggtgaag	gagagaattg	180
cattgtgttt	tgagaagttt	cactgacaag	ttatcctggg	ctgtgggaca	tcactagctt	240
tgaaagtgt	gctggcacct	cgtccatcta	atgtgatggg	tgtgtgtggg	gtgttgggca	300
cgcgtcggcc	tagcagatct	gaacccaggt	gatttctgtt	ctcaggaagc	ttttaggtga	360
caaggatcag	gcatgtgaac	aaataacct	actgg			395

<210> 119

<211> 144

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> (1)...(144)

<223> n = A,T,C or G

<400> 119

ccggttaagga	atatacttct	tctgatacta	aatatgccaa	tattttaa	gtaatattca	60
gggattacaa	ctgtgagggc	taaacacacg	gaattaccca	ccaattcctc	tgtagttctc	120
tactaattca	attttgcattc	ctcn				144

<210> 120

<211> 392

<212> DNA

<213> Homo sapiens

<400> 120

ggcacgagac	caggtcataa	gaggatccgt	tccaatgatt	ttcctaaaac	aatggaagt	60
ttttccaaag	agcttataag	gcattgtagg	atctggcctg	ccctgactcc	actttaccag	120
aaccatctgc	tgctcttctc	tcttgtgtta	ctcaaggtat	tagctgctgt	ggcaaataca	180
ctctgaaatc	tccgtgactt	aatacaagag	aggtttat	cttactcacg	ctgggtgcac	240
tgccacttgg	taacagagga	gctatggaaa	cttgagacct	aagcagaaat	gagttcaata	300
atattgctac	actctaggac	tttctccaaa	attaacaaca	gaacaaaagt	gcaaggcagt	360
gataacccat	ctgacagcat	ttggggagtg	tt			392

<210> 121

<211> 395

<212> DNA

<213> Homo sapiens

<400> 121

ggcacgagat	caatcacaaa	agttttatcct	taagacttcc	cttcagctgc	tggaaggcag	60
tcatcacatc	tgtgaaaaga	gtgctagtta	taacaaatga	gatcacaaat	ttgaccattt	120
tatttagacac	cctctatttag	tgtaacaga	caaagatgaa	ggtaagttg	aaatcaaatt	180
gaaatcatct	tccctctgta	cagattgcaa	tatctgataa	taccctcaac	tttcttggtg	240
caaattaatt	gcctgggtact	cacagtccag	tgtaaacagg	caataatggt	gtgattccag	300
aggagaggac	taggtggcag	gaaaataaat	gagattagca	gtatttgatt	ggagccataa	360
gcataatttg	gttccggcg	cggccaggtt	taaaa			395

<210> 122

<211> 288

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> (1)...(288)

<223> n = A,T,C or G

<400> 122

cgccccgcgc	tctctgttct	ctctcgcgcg	cggtgtctct	ctcgatagag	tgcgcgacct	60
gcacaccctc	tgtgtggggt	tctcgcctcc	cgtgtgcgcg	cgcgcgcgct	ctctgtggga	120

ctcgcacaca	ccgcgcgcgc	gcgcgcctctc	tgtggggggg	ccctccccgc	accttggtgtg	180
tgtgtgtctg	tggttatctct	gtgagatgtg	cgtgnnnnnn	nnnnntctgt	gtgtgtgtct	240
gccctccgcg	ccgtgtctgt	tatatatgcg	ctcgctcgct	ggggcgcg		288

<210> 123

<211> 393

<212> DNA

<213> Homo sapiens

<400> 123

ggcagcagga	tccattcttc	gacccccaga	tgtgactcta	aagaaggctg	aaaatttttg	60
tccaaattgc	catgcagata	tcttgaacag	caggacattt	gcaggccttg	tctactggac	120
ttttctccca	aacaggacaa	gccagggcag	ggctgcatgg	agagggaatg	aacctggagc	180
tagaattaat	tgcccactct	cccaccctac	cagtgcagcc	cggcaagggc	aggaattggg	240
aggcctaagg	tgggcatgaa	agcttgggaa	gcactgtcgt	ctctcagaca	ggcgtcctaa	300
agacctctag	gctggaagct	tgggcttgca	agtggatccg	ggaccgaggg	tggctctctg	360
gacaacccca	ggaacttgga	ccaaggcaga	gcc			393

<210> 124

<211> 394

<212> DNA

<213> Homo sapiens

<400> 124

ccgcgacgag	atgatgatct	gcttcttcca	ttatgccag	atgataaaaa	ggattgatac	60
aaaagaggga	tctgaattct	cattttcaga	tggagaagtg	gccgaaaaag	cagagggtta	120
caggtcagaa	aatgaaagt	aacggaactg	cctagaagaa	tcagagggct	gctattgcag	180
atcatctgga	gaccctgaac	aaataaagga	cgacagttaa	tcagaagaga	gtgctgatgc	240
acggagtttt	gaaatgactg	aactcaatca	agcttttaga	gaaataaaaag	ggcaggctgt	300
tgaaaacacc	tctgtaaactg	aattttctga	ggagaaacac	cgaacttgaa	attcacaccg	360
gcctaattgtc	caagaattca	aggggggggtc	cctc			394

<210> 125

<211> 390

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(390)

<223> n = A,T,C or G

<400> 125

ggcagcagcc	cttatacaaa	catatatgaa	catatatact	ttttttgttg	tataaaaaca	60
ggatcacatt	atagatatata	ttctgttaact	ttctgttttc	acccaaaata	cagcagagca	120
ctattttcca	gaagcacgta	gttctaactt	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	180
nnnnnnnnnn	nnnnnaactt	tattcaagta	cttcacattt	taagtggaca	ttccatttgt	240
ctgctataat	ttacaattat	agcaataactt	tgagaaagggt	ctttgcaagt	atatccatat	300
gaactaatgt	ctatgtagaa	gatatgctgg	ctcaaatatt	atgtacattt	aatgtcttaa	360
taaacaccgc	tagattactt	tccaggaagc				390

<210> 126

<211> 388

<212> DNA

<213> Homo sapiens

<400> 126

ggcagcaggt	cagcacacat	tactttaaca	ctttggactt	gaaattctga	aagatcagaa	60
attccttact	gtttgagatg	attaggtttt	agggactagc	cattttatct	cacatgactc	120
aggccttaat	gctccattgc	taatagctaa	atgtggaaaa	gtttagaatt	acattttaatt	180
tagtcaactg	ttaggctgca	atcatttttt	tttaaaaaatc	tgcttatggc	attattcgag	240
ataacttgac	caactctaaa	atatatatgt	aattacttct	agatgtaagt	agtttttcat	300
attaacaaca	caatcaggct	ctgttttcagt	tagttcttag	agtgggtgaaa	aaaaatcttt	360

acagtaagtg caaaattata atccaagg

388

<210> 127

<211> 388

<212> DNA

<213> Homo sapiens

<400> 127

ggcacgagag	ttaatccaaa	agacttcctt	tcagctgctg	gaaatcagtc	atcacatctg	60
tgaaaagagt	gctagttata	acaaatgaga	tcacaaattt	gaccatttta	ttagacaccc	120
tctattagtg	ttaacagaca	aagatgaagg	ttaagttgaa	atcaaattga	aatcatcttc	180
cctctgtaca	gattgcaata	tctgataata	ccctcaactt	tcttggtgca	aattaattgc	240
ctggtactca	cagtccagtg	ttaacaggca	ataatggtgt	gattccagag	gagaggacta	300
ggtggcagga	aaataaatga	gattagcagt	atttgacttg	gagccatagg	catcaattct	360
gctccagctg	tcgaccaggt	tctaaaaa				388

<210> 128

<211> 267

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(267)

<223> n = A,T,C or G

<400> 128

actgtgtgtg	tgtctgtttt	ctctctctct	cttctcagtc	acactttttt	tttgggacac	60
accctccatc	cgcggggggg	tttttttccc	ggcgcgcgcc	cttttttttt	gtgtgtttct	120
ctgcgcgcct	ctcttttttc	tctctcttcc	ccccccgctt	annnnnnnnn	nnnnnngcgg	180
gggggggttt	cgcgcgttcn	nnnnnnnnnn	nnctcttccg	cccccccaca	gggggggtgct	240
gtttattatc	tttctttctc	cctgagc				267

<210> 129

<211> 389

<212> DNA

<213> Homo sapiens

<400> 129

ggcacgagct	tgactgcaaa	cttgctgaag	gtagggactg	tttgtcttgg	acttcgctgc	60
cagtccttag	aacagtgtct	gggacacagt	gtgttctcaa	atatttggtg	ctggaataaa	120
tgaatgaact	aaatcagtct	tttagggatt	tactgttaac	caccatggga	aaattaaata	180
aatgcgggga	aggaaaacgt	tctaaaatta	gaagactact	ttctactctc	agcttctgat	240
tccctctgag	ctaagaacca	gacagcctta	ggctggtaac	tcctataagc	tggtcctcct	300
cccatgctga	ccccatcttt	actgtacaat	tcacttttca	tggactgaag	gcaccaccaa	360
gatagatcca	ggagtgacaa	ctccagtg				389

<210> 130

<211> 319

<212> DNA

<213> Homo sapiens

<400> 130

tggttgtaact	gggagtggag	gcccagtggc	tggggagaca	ttaggtggtg	gggcccagcc	60
cgacctccag	gttcttcctt	ctccctagct	gttgcttttg	tctggccact	cccagccccc	120
ttgtcccctt	ggaagcttgc	cctgccctca	tcttgcccat	gccttctact	gccaggagac	180
ttgcacccat	ttcaacccta	gggcgggggc	aagtggggca	aggatggacc	agcagaaggg	240
gggtaaggct	ctgttcactt	ccccctgcct	ccacagaacg	aagccacgga	ttccgttatc	300
ttcctccagt	tttgttcct					319

<210> 131

<211> 385

<212> DNA

<213> Homo sapiens

<400> 131

ggcacgagaa	acgttttcagc	tacgaaagtg	agctttttcc	aactttttaag	gatatcagga	60
gagaagacac	tcttgatgtg	gaggtttctg	ccagtggcta	cacaaaggaa	atgcaggcag	120
atgatgaact	gcttcatcca	ttaggtccag	atgataaaaa	tattgaaaca	aaagagggat	180
ctgaattctc	attttcagat	ggagaagtgg	cagaaaaagc	agaggtttac	aggtcagaaa	240
atgaaagtga	acggaactgt	ctagaagaat	cagagggtg	ctattgcaga	tcatctggag	300
accctgaaca	aataaaggaa	gacagtttat	cagaagagag	tgctgatgca	cggagttttg	360
aatgactga	attcaatcaa	gcttt				385

<210> 132

<211> 383

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(383)

<223> n = A,T,C or G

<400> 132

ggcacgaggg	gaatagaggg	tccctggtga	cagggcaagg	ctagatctgg	agcctgcact	60
tggcctgtga	catactgtct	tgtttctgag	aatcctcccc	tacttctcta	gataatctcc	120
aaacacttct	gtgactactt	aatcacaaag	gaaattttca	ggagatataa	tcgaattcta	180
ttttacaaaa	aaaaagagaa	gggatctgaa	tgttttcagt	tcacgctagg	gatcnnnnnn	240
nnnnnnnnnc	ccaaacctga	cgtttgagga	cccgcctttt	tttcagccaa	tttaaaagat	300
tttttaaggt	ttagggttgg	ttggccatta	aaccatcccc	ggaaagaaaa	tgggggtaaa	360
agaccaagaa	ggaggtcgcc	aag				383

<210> 133

<211> 382

<212> DNA

<213> Homo sapiens

<400> 133

ggcacgagat	aagatctgag	gtgttacaca	cataattgtc	ccaattttta	agattgatgg	60
ggagcatgaa	gcattttttt	aatgtgttgg	caggcccat	taaatgcata	aactgcatag	120
gactcatgtg	gtctgaatgt	attttagggc	tttctgggaa	ttgtcttgac	agagaacctc	180
agctggacaa	agcagccttg	atctgagtga	gctaactgac	acaatgaaac	tgtcaggcat	240
gtttctgctc	ctctctctgg	ctcttttctg	ctttttaaca	ggtgtcttca	gtcagggagg	300
acaggttgac	tgtggtgagt	tccaggacac	caaggtctac	tgcactcggg	aatctaacc	360
acactggggc	cttgaatggc	ca				382

<210> 134

<211> 375

<212> DNA

<213> Homo sapiens

<400> 134

ggcacgagca	agcctttcat	ttacattaaa	ttataacttt	tcattcattc	ctaaacccaa	60
cttaaaattc	tgcttttctt	tgagtagaag	gtattttaact	tgttttgttt	ttccttcaga	120
aggaatttaa	tgcaaacgga	ttgcagtcag	cacttttctga	atgttttcac	acagtatgca	180
aagcttacat	cataccaagg	agtggagagt	tgaagtttcc	tcccagtgac	tccagtgaca	240
gaccacacct	agaaagcgtt	tctcttctctg	agtatttcaa	aaagatgtaa	aagagctggg	300
gagagtatgg	gaagaaacaa	tacaggattg	cctttaatta	attaagaatt	gcctcctgat	360
aaaaggaaaa	agaaa					375

<210> 135

<211> 376

<212> DNA

<213> Homo sapiens

<400> 135
 ggcacgagac ctgtttgagg tggaactcca agcagctcgc accttggagc gactggagct 60
 ccagagtctg gaggcagctg agatagagcc ggaggcccag gccagagggt cggccaggcc 120
 cacgggctca gatctgctcc ctggagcccc catcctcagt ctgcgcttct cctacatctg 180
 ccctgaccgg cagttgcgtc gctatttggg gctggagcct gatgcccaag cagctgtcca 240
 ggagctgctt gccgtgttga cccagtcac caatgtggct gttcccctgc aggatctgag 300
 tggcatagag ctgggcctgg caggccagag cctgcggcta gagtgggcag ctggggcggg 360
 ccgctgtgtg ctgctg 376

<210> 136
 <211> 371
 <212> DNA
 <213> Homo sapiens

<400> 136
 ggcacgaggt cacctctacc agccctcctt tctccagatg gcttcttcat aaccaccagg 60
 tcagaagagg atccgttcca atgattttcc taaaacaatg gaagtgtttt ccaaagagct 120
 tataaggcat tgtaggatct ggcctgccct gactccactt taccagaacc atctgctgct 180
 cttctctctt gtgttactca aggtattagc tgctgtggca aatcaactct gaaatctccg 240
 tgacttaata caagagagg ttatttctta ctacgctgg gtgcactgcc acttggtaac 300
 agaggagcta tggaaacttg agacctaagc agaaatgagt tcaataatat tgctacactc 360
 taggactttc t 371

<210> 137
 <211> 258
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(258)
 <223> n = A,T,C or G

<400> 137
 cagtttcttt gtgcgcgcgc cccccctttt ttctctctct ctccgcgcgg gcgtgtccct 60
 ccnnnnnnnn nnctgtgtgt gcgctctctc cgccccatat atattgtgtt tttctctgtg 120
 gannnnnnnn nntctctcta gagtcttttc tctcccctcg cgcgacatt gttatacact 180
 cctcccctct ctttcttttt acacacacat atatattgcy cccctctccc cccacacatt 240
 tatactctct tcacatct 258

<210> 138
 <211> 368
 <212> DNA
 <213> Homo sapiens

<400> 138
 ggcacgagac attttgagac ttcttccaaa ttggtcccta gaaagttaca ctggtttgta 60
 ctctcactta tgtcactgtt tataccacca ctgactgctg cctgctttat tatttcttta 120
 atgagttgga ctgaacagtg gttaatcctg actctgtttt tgactgacag ttaacagtta 180
 catgaaccat tcatattaca gctcttactt aaatttgacc aagccaggat atatctgtta 240
 ggccacattc atttagggat catgttttcc aaagcaggtt tgggcaaaat taatccacag 300
 gactgaaagg tatacatctg tgagttttgt tctcacttcc acctctaatt tgaagaacac 360
 tttaattg 368

<210> 139
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 139
 acggcacgag ctggctcctc gttttctttt tggacagtct cattaccaac atcctcgttc 60
 gggcttagga tgcctttctg ctcgagggga ccaacgcggc gattcgctat gccttggcca 120
 ttatcttgta caacgagaag gacatcttga ggctacagaa tggcctggaa atctaccagg 180

acctgcgctt	cttcaccaat	accaactcca	tcagccggaa	gctgatgaac	attgccttca	240
atgacatgaa	ccccctccgc	atgaaactat	tgccggcagct	gtgcatggcc	caccgtgagc	300
ggctggaggc	tgatctgccg	gagctggagc	aacttaaggc	aaagtacctg	gctaggcagg	360
catcccggcg	ca					372

<210> 140
 <211> 365
 <212> DNA
 <213> Homo sapiens

<400> 140						
ggcacgaggc	tgagagtgtc	tgatacacc	ttgaatcccc	tcttatatga	tgccccagcc	60
caggagagat	aaaagcatca	gcacatgag	attcacctgc	ctctggctcg	tagggaacaa	120
tggaggcctg	cgatttgagg	ttaaactctc	agtgatctct	gtgttgacaa	caccaaagct	180
agaggaatcc	agtaggatgt	gggcatgggt	ttcccgggaag	gctgactgag	cagttctgca	240
aatgttttgc	agtacagggc	agaatttcat	ccagcctcag	aaccttgagc	caagactcag	300
catcagcaaa	gccaaaagtt	tcatttcttt	gactgtggga	gtgctagtcc	caacctttag	360
atggc						365

<210> 141
 <211> 353
 <212> DNA
 <213> Homo sapiens

<400> 141						
ggcacgagaa	acaaaagaga	gcaagagaga	agacagtggg	tgaagtcctg	gttccagact	60
cccctttttg	ccgggatatg	atggatctgt	cagctggtga	ggccctctta	agaggggtgg	120
tatcttcggg	ccaggtgcct	agagtcctag	agagctagag	atggaggga	attcagatca	180
tctaaaccct	tcagcccttc	actggacaga	agaggaaact	gaggctccat	ctgcatgacg	240
ttcccagagt	cacggcacaa	attcatggaa	gaagcagcag	gaaactcagt	tctccagtct	300
gggtccaatg	tgtgttttag	aaatatctcc	acagggttaa	tgactcaatt	ttt	353

<210> 142
 <211> 352
 <212> DNA
 <213> Homo sapiens

<400> 142						
ggcacgaggc	cactcggggg	cccaggaacc	cctcagttag	ggcttctcag	tactgagcgc	60
gaaggtgccc	ccagaggggg	cagccgcctg	tgaggagcag	gcgtgtctgg	gtaaccatgt	120
ggctcctgct	ggcctcccc	gctgtcccc	aaagcacagg	gctcagctcc	agagggagac	180
gggctgggct	gtcagtggtc	ccaggtgcat	cccactttcc	agcagcactt	ggtgccagca	240
gaggctgcag	gtgtggcagg	agggggccca	gccgtgaggg	caccagggtc	aggcccgga	300
tctcaggggtg	gagagccagg	gctgtcctga	acctccagag	ggggtgagct	gg	352

<210> 143
 <211> 470
 <212> DNA
 <213> Homo sapiens

<400> 143						
gacttctgtc	tttttaggat	cccatcgact	tcaattcggc	acgaggtcat	gagaaaggaa	60
ccaatggagt	atgagaagtt	tccagtga	aacagaaaga	atccagtaga	atttattttag	120
ggaagaggaa	aagatgtgtt	cgggggtggc	ttggaagtga	acgttgaagg	actactgaga	180
ttggttcaag	aaactgtgaa	gggaaagaaa	gggttatact	gagaaatgga	agagataatt	240
ttagaaactt	gcgaaaaatg	gcttaatcta	aatgagtgtt	aggggagata	cagctgtgat	300
gataggttga	gctcacatgg	tggagagcca	cagttgcggg	tgcttgcact	gataatgtga	360
gggcatggag	acagacaata	agttgaatgc	tcttttttta	acaaagggaag	ctaaaaggga	420
gggggatgct	aatttgatca	atacgttttg	gaaaacttat	atcttcttgg		470

<210> 144
 <211> 456
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(456)

<223> n = A,T,C or G

<400> 144

tagcactttt	gtttaggagg	accccatcga	ttcgaattcg	gcacgagctg	cactgagcag	60
caccggtggt	cttcatccgg	ctgcaccccc	aacagagctc	tttcttcccc	agatcccttt	120
tacagttgga	ttctccctct	tggatctggc	tctgccttag	tccgacctag	agggatcagc	180
ttcgcccacg	cccactctca	cccggaacct	ttcatctctt	attgaagcct	tttaggcccc	240
ttgggatggt	cattagaact	ctgaaaacta	cagttctccc	ctttatgagg	actgcaccac	300
agctcgccct	ctcctgggtt	ccgcctgggt	gcagagttag	cccatgggac	agccctctga	360
aattatactg	cttacaacca	tgctgagtct	gcaaggactt	cgtccaagcc	tttccgtcca	420
ggacctcaaa	cagatccaat	cacaagaaga	gagatn			456

<210> 145

<211> 464

<212> DNA

<213> Homo sapiens

<400> 145

atcgcccata	cggcgagccc	accgacgcga	attcggcacg	aggggaaaca	caggcctctt	60
ctgcttttag	gaccctcccc	ctgccttgca	gggggctcgg	ggagagcaat	atcaggagct	120
agggccttgc	gctgcccaca	ctcctgcttt	ttgggatata	taactgctaa	ggagggagtt	180
gacatccccc	ttctgggtca	tgtgtctgac	accaacaaca	tgggctctgt	ccctctctct	240
ttgactctcc	ctttgtcctc	cccatacagc	tggggtgggg	tggatcccta	tacctggggc	300
aggcagcccc	aaagtgggtg	agggggatgg	caaagactgt	ataggcgcca	ctggactctg	360
gcaaggcctt	tattaccttt	actccccttc	ctctcccatc	accagcctca	aggcctgagg	420
tgtgcagggg	ctcctggcag	ctactgagtg	agggttccctg	gtcg		464

<210> 146

<211> 448

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(448)

<223> n = A,T,C or G

<400> 146

ggcagagct	gcactgagca	gcaccggtgt	tcttcatccg	gctgcacccc	caacagagct	60
ctttcttccc	cagatccctt	ttacagttgg	attctccctc	ttggatctgg	ctctgcctta	120
gtccgacct	gagggatcag	cttcgcccac	gccactctc	acccggaacc	tttcatctct	180
tattgaagcc	ttttaggccc	attgggatgt	tcattagaac	tctgaaaact	acagttctcc	240
cctttatgag	gactgcacca	cagctcgccc	tctcctgggt	tccgcctggg	tgcagagtga	300
gcccatggga	cagccctctg	aaattatact	gcttacaacc	atgctgagtc	tgcaaggact	360
tcgtccaagc	ctttccgtcc	aggacctcaa	acagatccaa	tcacaagaag	agagatttca	420
ggaaagagaa	nattattcct	atcatcgn				448

<210> 147

<211> 439

<212> DNA

<213> Homo sapiens

<400> 147

ggcagagga	aagttaagca	actacaggaa	atggcttttg	gagttccaat	atcagtctat	60
cttttattca	acgcaatgac	agcactgacc	gaagaggcag	ccgtgactgt	aacacctcca	120
atcacagccc	agcaagctga	caacatagaa	ggacccatag	ccttgaagtt	ctcacacctt	180
tgcttggaag	atcataacag	ttactgcata	aacggtgctt	gtgcattcca	ccatgagcta	240
gagaaagcca	tctgcaggtg	ttttactggg	tatactggag	aaaggtgtga	gcacttgact	300

ttaaacttcat	atgctgtgga	ttcttatgaa	aaatacattg	caattgggat	tggtgttgga	360
ttactatttaa	gtgggttttct	tgttatTTTT	tactgctata	taagaaagag	gtgtctaaaa	420
ttgaaatcgc	cttacaatg					439

<210> 148
 <211> 334
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(334)
 <223> n = A,T,C or G

<400> 148						
ccccgcgcgc	gctccctctc	tatcttttat	acaaaatata	gagagcgcac	atctctgtgt	60
gtgagagagt	ctgtgcgcgc	gcgcataatat	atatgggagg	ggtgtctccc	cccatctgtg	120
tgtctctcct	cttgcggggc	atatgcgtgc	gcacacccgc	gcgctgtgtc	tcttttgtgc	180
cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	300
gtgtgttcta	cagcgcgata	aagagagaca	caca			334

<210> 149
 <211> 428
 <212> DNA
 <213> Homo sapiens

<400> 149						
ggcagcaggt	cctgagcagc	ctcatgggag	gtgaattaga	gaaaacaaaa	gagagcaaga	60
gagaagacag	tgggtgaagt	cctgggtcca	gactcccctt	tttgccggga	tatgatggat	120
ctgtcagctg	gtgcctagag	tcctagagag	ctagagatgg	agggaaattc	agatcatcta	180
aacccttcag	cccttcactg	gacagaagag	gaaactgagg	ctccatctgc	atgacgttcc	240
cagagtcacg	gcacaaattc	atggaagaag	cagcaggaaa	ctcagttctc	cagtctgggt	300
ccaatgtgtg	tttttagaaat	atctccacag	ggttaatgac	tcaatttttc	atgcatgatt	360
gctagtaatg	acaatcatgt	tatgttttgt	tctgtagctt	tggaatcac	tccttccact	420
tgagtttc						428

<210> 150
 <211> 427
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(427)
 <223> n = A,T,C or G

<400> 150						
cgccccaaan	nnnaatctct	aaaggggtaa	gggagatacc	taccttgtct	ggtaggggag	60
atgtttcggt	ttcatgcttt	accagaaaat	ccacttccct	gccgacctta	gtttcaaagc	120
ttattcttaa	ttagagacaa	gaaacctgtt	tcaacttgaa	gacaccgtat	gaggtgaatg	180
gacagccagc	caccacaatg	aaagaaatca	aaccaggaaat	aacctatgct	gaaccacgc	240
ctcaatcgtc	ccaagtgtt	tcctgacacg	catctttgct	tacagtgcac	cacaactgaa	300
gaatggggtt	caacttgacg	cttgcaaaat	taccaaataa	cgagctgcac	ggccaagaga	360
gtcacaattc	aggcaacagc	agcgacgggc	caggaaagaa	caccaccctt	cacaatgaat	420
ttgacac						427

<210> 151
 <211> 437
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature
 <222> (1)...(437)
 <223> n = A,T,C or G

<400> 151
 ccgagccgga tgnccctnnn gagtatngca angattccaa ttcggcacga gagacagtgg 60
 catggagctt tgaaagacga gtaggtgtta gcaaggaaat aaggaggaac ggggggttacg 120
 ggcagaggag aaagcacatg ccaagtcagc aaagaaaagt agaattcgaa aacttttttaa 180
 aaatattact aaggattttc acaatgctgc actgggctag aaactgaagc taaaacagat 240
 acgtgggtccc tgctgctatg gggcttacgt tctacaggca aggacagggt gtgatgaggg 300
 ttctgaagga tagagaccaa gcatggaggg tggtgaggag gcttctgcga gacctgaatg 360
 atgggaagcc acgaagtggg aggggtgggg gtccaggctg gaggggcccc atgtatgtgt 420
 agagggacta cagccct 437

<210> 152
 <211> 425
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(425)
 <223> n = A,T,C or G

<400> 152
 ggcacgagct gcaactgagca gcaccgggtg tcttcattccg gctgcacccc caacagagct 60
 ctttcttccc cagatccctt ttacagttgg attctccctc ttggatctgg ctctgcctta 120
 gtccgacctt gagggatcag cttcgcccac gccactctc acccggaacc tttcatctct 180
 tattgaagcc ttttaggccc attgggatgt tcattagaac tctgaaaact acagttctcc 240
 cttttatgag gactgcacca cagctcgccc tctcctgggt tccgcctggg tgcagagtga 300
 gcccatggga cagccctctg aaattatact gcttacaacc atgctgagtc tgcaaggact 360
 tcgtccaagc ctttccgtcc agggacctca acagatccaa tcacaagaag agagatttca 420
 ggaan 425

<210> 153
 <211> 421
 <212> DNA
 <213> Homo sapiens

<400> 153
 ggcacgagcc gtggctgcct cgtgagcctc ccagagccca ggccctccgtg gcctcctcct 60
 gtgtgagtc caccaggagc cacgtgcccg gccttgccct caaggatttt tgcttttctc 120
 ctgtgcacct ggcgaggctg aaggcgaggg gtggaggagg cccagacaca gcctcatctc 180
 catgtgtaca cgtgtgtacg tgtgtatgcg tgtgtgtacg tgtgtatgcg tgtgtgtacg 240
 cgtgtgtacg tgcgtgtgta cacatgcgtg gccgcctgtg gtgtgcacgt gtgctctggg 300
 ctccgaggct tctccagagc tgggagctgg ctggcgtggc aagggcacgc tctggggcag 360
 tgtgtccctc aggaaccagg gtcctccctc ccctttctgc ctggtcagcc ccgtggcctc 420
 t 421

<210> 154
 <211> 423
 <212> DNA
 <213> Homo sapiens

<400> 154
 ggcagtagtg gaagggaggg agctgccttt gtttgccatg gatgggtagg ggctgcactg 60
 agcagcagcg gtgttcttca tccggctgca cccccaacag agctctttct tccccagatc 120
 cctttttacag ttggattctc cctcttggat ctggctctgc cttagtccga cctagaggga 180
 tcagcttcgc ccacgcccac tctcaccggg aacctttcat ctcttattga agccttttag 240
 gccattggg atgttcatta gaactctgaa aactacagtt ctccccctta tgaggactgc 300
 accacagctc gccctctcct gggttccgcc tggttgcaga gtgagcccat gggacagccc 360
 tctgaaatta tactgcttac aacctgctg agtctgcaag gacttccgcc aagcctttcc 420
 gtc 423

<210> 155
 <211> 312
 <212> DNA
 <213> Homo sapiens

<400> 155
 tctgtcactc acaaaacaca gtgcgcgcac atagcggggg gggagcacac acacaagatg 60
 tgttgtgtata caaccgcgc gcgagagagc gctctctttt gtggggggga aaaaaactct 120
 tatacacaca cgtgtgtgtg tgtcgcctctc cgaaaataca cactataaca aacgcactgt 180
 gtgtgtgaga cacacactcc tctctccgag tggggagaga gagatcgcg cccactctta 240
 aacacatatg cgctcacaga gagcatatat atgttttttt tgagagaaga gagagatctc 300
 tttgtggttt ct 312

<210> 156
 <211> 428
 <212> DNA
 <213> Homo sapiens

<400> 156
 tgaccttcca ggctacctac gcaggtgtcg gggccaacaa gcacctgcag gagctggccc 60
 aggaggaggt gaagcagcat gccaggaac tctgggctgc ctacaggggt ctgctgcgag 120
 ttgccttaga gcgcaagggc caggccctgg aggaggatga agacacagag acaagggacc 180
 tccaggtgca tggatttgtg ctgcccctca tgctgcccag cttctactca gagctcttca 240
 cgctctacct gctgcttcat gagcgggagg acagcttcta cagccagggc attgccaact 300
 tgagcctctt tcctgatacc caactgctcg agttcctgga tgtgcagaag cacttggtggc 360
 cctcaagga cctcacgctg acgagcaatc agaggtactc cctgggtcagg gacaagtgtt 420
 tcctgtca 428

<210> 157
 <211> 430
 <212> DNA
 <213> Homo sapiens

<400> 157
 ggcacgagag gactttgagc ccagagagat gaagtcattt gctcaaggca gcagtcagtg 60
 gaagggcttg gagaaggaga aggggtctga aggtggtgtg ggacacatga gagtgatctc 120
 gcagcttggt ttgctgcagc agactcggac aagcattgtt tcagtgcctg gtttctccct 180
 ccacttgatg ggggccaact ccaaccatgt gtcccattcc tatcctgaaa tgcttctaaa 240
 ggcagtgcc tgagaaccac caacctcaca gcctgtctcc attttattgt cttctgggaa 300
 cttctccctt ctgtctagca cctgtttgca ctgggattgt cctgtctgtc cttcagttgg 360
 atcctggttt gcacccgatg aggatttagc aatttttaggc tgtgcttcgg caaaggccaa 420
 ctcacaatgg 430

<210> 158
 <211> 405
 <212> DNA
 <213> Homo sapiens

<400> 158
 ggcacgaggg aagatttcca gtggtctcaa tgggtgtgaat cctatgaagg tgtcttattt 60
 gttgaattag aggtgaaagc ctcttctctc actctttttt agaaacagtt tagttttatt 120
 attatgcaga atttggtgag caaattgcaa cagcccaagc cacagctagc tccacaagag 180
 cccttccatg agccctcaac ctgggatctc gtgtatcttt gttggaatgg acattaggtt 240
 tccaagtcca ggcctgtgat ttagaagggt caggttgggt aggagagagg agagtcttgg 300
 aggggctgct ccatgggggt cacacctctc tcctgtgggt tttcgctggg gattgagttc 360
 tgaggcattt gctgcattga ctgtttagc ttttaactcg gtgca 405

<210> 159
 <211> 403
 <212> DNA
 <213> Homo sapiens

<400> 159

```

ggcacgagcc tgactcaagg ggttttggaa gatttccagt ggtctcaatg gtgtgaatcc      60
tatgaagggtg tcttatttgt tgaattagag gtgaaagcct ccttcctcac tcttttttag      120
aaacagtttta gttttattat tatgcagaat ttgttgagca aattgcaaca gccaagcca      180
cagctagctc cacaagagcc cttccatgag ccctcaacct gggatctcgt gtatctttgt      240
tggaatggac attaggtttc caagtccagg cctgtgattt agaaggggtca ggttgggtag      300
gagagaggag agtcttggag gggctgctcc atgggggtca cacctctctc ctgtgggttt      360
tcgctgggtga ttgagttctg aggcatttgc tgcattgact gtg                        403

```

<210> 160

<211> 417

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(417)

<223> n = A,T,C or G

<400> 160

```

gttctgtggg aatagagggt ccctggtgac agggcagggc tagatctgga gcctgcactt      60
ggcctgtgac atactgtctt gtttctgaga atcctccctt acttctctag ttaatctcca      120
gagacttctg tgactactta atcacaaagg aaattttcag gaatattatc aaatactatt      180
ttagaaaaaa aaagagaagg gatttgaatg ttttcagttc agtttagtta tcnnnnnnnn      240
nnnnnnnccc caaactccag aatggggggc ccccttctt taacccacc taaaaatttt      300
tcggagggttc agggttggtt ggcaaattac aaaaacccca aaagaaaatg ggggttaacc      360
cccttggaag agttttctta ctttgggggg tggccctttg acgtnggccg ggggttac      417

```

<210> 161

<211> 300

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(300)

<223> n = A,T,C or G

<400> 161

```

ctatatctct ctgcgccctc tccccctctt gtgttttccc ccgccctct agagatatct      60
ctctcactcg cgggcgcaca ccccccttta caaaataggg ggctctctgt gtgtgggtgt      120
tttcttgggc gccccctctt tttttttctt tttgcggg ccgccctgtg tgtctctctc      180
tagacacacc ccccgcgcg tgttttttat aaatatctgt ctctcacaca cccctactg      240
cccctctgtg tgtgggcgcg ttcccccca cacacacaga gtgtgtgnnn nnnnnnnnnn      300

```

<210> 162

<211> 411

<212> DNA

<213> Homo sapiens

<400> 162

```

ggcacgaggg caccgagcct cctgtgggag gtcccagagg agcttcgcct gctcggcctg      60
gctgcagccc tcacctgccg cagccttagc tgagcagccg ccgccactgg gcgccccccg      120
ctccccactt cgccagcgcc cgctcctcgg ctcggcccg ggtagtttgt agggacgcag      180
ctctccacgt gcgcgactgc gaggtggac gctacgggct cctggaaagg agcagacacc      240
agcatttgcc acaatgctgt catccactga ctttacattt gcttcctggg agcttgtggt      300
ccgcgttgac catcccaatg aagagcaggc agaaagacgt ccgcactgag aggattctgg      360
agacccttca cgttggaagg agtgatgctc aaggtttagta gaacagatca a                        411

```

<210> 163

<211> 412

<212> DNA

<213> Homo sapiens

```

<400> 163
gcacgatcca tcattcaatt cggacagcca ctccaactga cctgttccgt ggctgcctcg      60
agagcctccc atagcccagg cctccgtagg cctcctcctg tgtgagtcct accaggagcc      120
acgtgcccgg ccttgccctc aagggttttt gcttttctcc tgtgcacctg gctaggctga      180
aggcgagggg tggaggaggc cccagcacag cctcatctcc atgtgtacac gtgtgtacgt      240
gtgtatgcgt gtgtgtacgc gtgtacgcgt gtgtgtacgc gtgagtacgt gctgtgtgta      300
cacatgcgtg gccgcctgtg gtgtgcacgt gtgctctggg ctccgaggct tctccagagc      360
tgggagctgg ctggcgtggc aagggcacgc tctggggcag tgtgtccctc ag          412

```

```

<210> 164
<211> 411
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(411)
<223> n = A,T,C or G

```

```

<400> 164
ggcacgagag gatatggtgc aaaaaaatat gattttgtta accacaacaa aaagaaaggt      60
aagaaatgct aggagaaagc taaaagctcc atactaaaat aatggtccta atattaagca      120
aagtaaaatg tggatatgatt ttgagtgggc agcagagtgt aagaataatc tatttgcact      180
tgatactttc agctgtcaca gaggtcatag aattgggctt attgagaagg aaaggtaaat      240
gctagtacac tacttggctc agaagtgaac aaaattgcag tttgnnnnnn nnnnnnnnnn      300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn n          411

```

```

<210> 165
<211> 415
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(415)
<223> n = A,T,C or G

```

```

<400> 165
ggcacgagag gatatggtgc aaaaaaatat gattttgtta accacaacaa aaagaaaggt      60
aagaaatgct aggagaaagc taaaagctcc atactaaaat aatggtccta atattaagca      120
aagtaaaatg tggatatgatt ttgagtgggc agcagagtgt aagaataatc tatttgcact      180
tgatactttc agctgtcaca gaggtcatag aattgggctt attgagaagg aaaggtaaat      240
gctagtacac tacttggctc agaagtgaac aaaattgcag tttgnnnnnn nnnnnnnnnn      300
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn      360
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nngtn          415

```

```

<210> 166
<211> 403
<212> DNA
<213> Homo sapiens

```

```

<400> 166
ggcacgagga aggtgtcagg agcatcccat ttgtgtctct ctctctacct ctgtgaaggg      60
cgcgaaatgg cagagcagaa cttctagaag ggaagatgag caccaggat ccctcagatc      120
tgtggagcag atccgatgga gaggtcgagc tgctccagga cttgggggtgg tatcacggca      180
acctcacacg ccatgctgct gaagctcttc tcctctcaaa tggatgtgac ggcagctacc      240
ttctgagggg cagcaatgag accaccgggc tgtactctct ctctgtgagg gccaaagatt      300
ctgttaaaca ctttcatgtt gaatatactg gatattcatt taaatttggc tttaatgaat      360
tctcatcttt gaaggatttt gccaaagcatt ttgcaaatca gcg          403

```

```

<210> 167
<211> 407

```

<212> DNA
 <213> Homo sapiens

<400> 167
 ggcacgaggg ggcacaagct gttggagctg caatgggccc cggctgggga ttcttgtttg 60
 gcctcctggg cgcctgtgtg ctgctcagct cggggcacgg agaggagcag cccccggaga 120
 cagcggcaca gaggtgcttc tgccagggtta gtggttactt ggatgattgt acctgtgatg 180
 ttgaaacat tgatagattt aataactaca ggcttttccc aagactacaa aaacttcttg 240
 aaagtgacta ctttaggtat tacaaggtaa acctgaagag gccgtgtcct ttctggaatg 300
 acatcagcca gtgtggaaga agggactgtg ctgtcaaacc atgtcaatct gatgaagtgc 360
 ctgatggaat taaatctgcg agctacaagt attctgaaga agccaat 407

<210> 168
 <211> 416
 <212> DNA
 <213> Homo sapiens

<400> 168
 ggcacgagac acaactttga gacaccccaa gtgctttctg cagaggttgt cgttggaata 60
 ctgtcacctt acagaagcca attgcaagga ccttgctgct gtgttggttg tcagccggga 120
 gctgacacac ctgtgcttgg ccaagaaccc cattgggaat acaggggtga agtttctgtg 180
 tgagggcttg aggtacccc agtgtaaaact gcagaccttg gtgctttgga actgacacat 240
 aactagcgat ggctgctgcg atctcacaaa gcttctccaa gaaaaatcaa gcctgttgtg 300
 tttggatctg gggctgaatc acataggagt taagggaatg aagttcctgt gtgaggcttt 360
 gaggaacca ctgtgcaact tgagatgtct gtggttgttg ggatgttcca tccctc 416

<210> 169
 <211> 386
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(386)
 <223> n = A,T,C or G

<400> 169
 ggcacgagga atctcgctc tgtctggtgt gttacctact gggggcacag gaacaatttc 60
 ctcaaggaga cagtggcatg gagctttgaa agacgagtag gtgttagcaa ggaaataagg 120
 aggaacgggg gttacgggca gaggagaaag cacatgccaa gtcagcaaag aaaagtagaa 180
 ttcgaaaact ttttaaaaat attactaagg attttcacaa tgctgcactg ggctagaaac 240
 tgaagctaaa acagatacgt ggtccctgct gctatggggc ttccgttcta gaggcaagga 300
 caggttgtga tgaggggttct gaaggataga gaccaagcag ggaggggtgt gaggaggctt 360
 ctgcgagacc tgaaggatgg gaagcn 386

<210> 170
 <211> 391
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(391)
 <223> n = A,T,C or G

<400> 170
 ggcacgagaa tagagggtcc ctggtgacag ggcagggcta gatctggagc ctgcacttgg 60
 cctgtgacat actgtcttgt ttctgagaat cctcccctac ttctctagtt aatctccaga 120
 gacttctgtg actacttaat cacaaggaa attttcagga atattatcaa atactatatt 180
 agaaaaaaaa agagaaggga tttgaatgtt tttagttcag ttttagttatc nnnnnnnnnn 240
 nnnnncccaa aactcaagat tggggccccc ccctccttta accccgctaa aaagtttttt 300
 gggggtttag ggtgggttgg caaataacaa aacccccaaa agaaaagggg ggtaaaccct 360
 cttggaaaag tttcctaact ttggggggcg c 391

<210> 171
 <211> 391
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(391)
 <223> n = A,T,C or G

<400> 171
 ggcacgagcc tgcacgacc catttttccct catgacaaac tattggtgca nnnnnnnnnn 60
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnact tagggccact 120
 catctgtcat ggaaccagaa tctaaatcca aataggctgt tgccagtaca gatggtaagt 180
 acatgtactt ctggcaggaa agcagaataa aagttgactg aacctgaaag tctcggaat 240
 ggtcttctca tttctattct gtaaagtgtc acgtcttcta ggcctacctc tgtcaatatt 300
 gaaatacaaa attaaactttt tctgcttttt atttcacaaa tcaacgggaa cagtcttagt 360
 catttgtgtt ttatgagttt taattaggcc n 391

<210> 172
 <211> 385
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(385)
 <223> n = A,T,C or G

<400> 172
 ggcacgagga cagtggcatg gagctttgaa agacgagtag gtgttagcaa ggaaataagg 60
 aggaacgggg gttacgggca gaggagaaag cacatgccaa gtcagcaaag aaaagtagaa 120
 ttcgaaaact ttttaaaaat attactaagg attttcacaa tgctgcactg ggctagaaac 180
 tgaagctaaa acagatacgt ggtccctgct gctatggggc ttccgttcta gaggcaagga 240
 caggttgtga tgagggttct gaaggataga gaccaagcag ggagggtgtt gaggaggctt 300
 ctgcgagacc tgaaggatgg gaagccagga agtgggaggg gtgggggtnc aggctggagg 360
 ggcccaatgt angtgtaaag ggact 385

<210> 173
 <211> 392
 <212> DNA
 <213> Homo sapiens

<400> 173
 ggcacgagaa aggctggaag ggaggcagct gcctttgttt gccatggatg ggtaggggct 60
 gcaactgagca gcaccggtgt tcttcatccg gctgcacccc cgacagagct ctttcttccc 120
 cagatccctt ttacagttgg attctccctc ttggatctgg ctctgcctta gtccgacctt 180
 gagggatcag cttcgccac gccactctc acccggaacc tttcatctct tattgaagcc 240
 ttttagggcc attgggatgt tcattagaac tctgaaaact acagttctcc cctttatgag 300
 gactgcacca cagctcgccc tctcctgggt tccgcctggt tgcagagtga gcccatggga 360
 cagccctctg aaattatact gttacaacc at 392

<210> 174
 <211> 394
 <212> DNA
 <213> Homo sapiens

<400> 174
 ggcacgagat ggaatgacag ctttttttag tagcatatcc ttgcgctgtg ttagatggag 60
 tctttgccct gatttccgtc ttttgaaaat ttatctggga tgtggacatc agtgggccag 120
 atgtacaaaa aggaccttga actcttaaat tggaccagca aactgctgca gcgcaactct 180
 catgcagatt tacatttgac tgttgagca atgaaagtaa acgtgtatct cttgttcatt 240
 tttatagaac ttttgcatac tatattggat ttacctgcgg tgtgactagc tttaaatggt 300

tgtgtttata	cagataagaa	atgctatttc	tttctgggtc	ctgcagccat	tggaaaaact	360
tttttctttg	gaaataataa	ggtttttgat	agat			394

<210> 175

<211> 387

<212> DNA

<213> Homo sapiens

<400> 175

ggcagcaggg	cagtttaggg	tgccatgtgc	tgggagctgt	gtgtctgctc	tccttcgtcc	60
gctccccag	ggcagtgtgg	tagcacatcc	cattgtagag	atgagggcac	cgaggcttcc	120
tggagcatac	cacctgggtcc	cgttcatgag	tggtaggcaa	gctagcactc	tcacttgtcc	180
attctgcctt	cctggagacc	agtgggatgg	gtcagtacag	cccaccacac	cattagcccc	240
aggaacataa	ggctgtggct	agacagcagg	ggtctcaggt	tcatacatga	ggactggctt	300
gtccttgagc	accactcac	ctgtctatgt	ggggaggaat	cctacaatag	gtcaccatgg	360
caggctgggg	cttgctgacc	ctgcccc				387

<210> 176

<211> 395

<212> DNA

<213> Homo sapiens

<400> 176

ggcagcagca	gacctccatt	acctccatcc	ctgttggatt	atttaaagaa	agcctcagac	60
agtaagggtc	ttttttaaaa	gaataaaatg	acttggtttg	cgcttggaag	caggggaagc	120
attcagatga	gcgggttctg	cattaaccct	gcctatcacg	catctcgtgt	cctgtgtggc	180
tggcgagccc	cccttggaag	gttctgggtc	ttcagctggc	tcctgcagag	tcacccccgc	240
ctcgtgggtg	gaatgcagag	ccctttgctt	tccttcttgc	cgctgcttc	ctgttctctg	300
ggacccgctg	ggccttttgt	ctgcatcccc	tggccaggtc	cctcagggtc	gatgcgcgta	360
gaaggacttt	gagcagtggg	ggcagcactt	gcct			395

<210> 177

<211> 388

<212> DNA

<213> Homo sapiens

<400> 177

ggcagcaggg	acgctgcgga	gcccgcctcac	ccgctccctg	tacgtgaaca	tgactagcgg	60
cccgggtggg	ccggcgggcg	ccgcgggcg	caggaaggag	aaccaccagt	ggtatgtgtg	120
caacagagag	aaattatgcg	aatcactcca	ggctgtcttt	gttcagagtt	accttgatca	180
aggaacacag	atcttcttaa	acaacagcat	tgagaaatcg	ggctggctat	ttatccaatt	240
atatcattct	tttgtgtcat	ctgttttttag	cctgtttatg	tctagaacat	ctatcaatgg	300
gttgctagga	agagggtcaa	tgtttgtgtt	ttcaccagat	cagtttcaga	gactgcttaa	360
aattaatcca	gactggaaaa	cccacaga				388

<210> 178

<211> 397

<212> DNA

<213> Homo sapiens

<400> 178

ggcagcagca	ggatccctca	gatctgtgga	gcagatccga	tggagaggct	gagctgctcc	60
aggacttggg	gtggtatcac	ggcaacctca	cacgccatgc	tgctgaagct	cttctcctct	120
caaattgatg	tgacggcagc	taccttctga	gggacagcaa	tgagaccacc	gggctgtact	180
ctctctctgt	gagggccaaa	gattctgtta	aacactttca	tgttgaatat	actggatatt	240
catttaaatt	tggctttaat	gaattctcat	ctttgaagga	ttttgtcaag	cattttgcaa	300
atcagccttt	gattggaagc	gagacaggca	ctctgatggg	tctaaaacat	ccctacccaa	360
gaaaagtgga	agaaccctcc	atttatgaat	ctgtccg			397

<210> 179

<211> 397

<212> DNA

<213> Homo sapiens

```

<400> 179
ggcacgaggc gtggggcgac aagctgccgg agctgcaatg ggccgcggct ggggattctt      60
gtttggcctc ctgggcgccg tgtggctgct cagctcgggc cacggagagg agcagcccc      120
ggagacagcg gcacagaggt gcttctgccca ggtagtggt tacttggatg attgtacctg      180
tgatgttgaa accattgata gatttaataa ctacaggctt ttccaagac tacaaaaact      240
tcttgaagtg gactacttta ggtattacaa ggtaaacctg aagaggccgt gtcctttctg      300
gaatgacatc agccagtgtg gaagaaggga ctgtgctgtc aaacctatgtc aatctgatga      360
agttcctgat ggaattaaat ctgcgagcta caagtat      397

```

```

<210> 180
<211> 399
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(399)
<223> n = A,T,C or G

```

```

<400> 180
ggcacgaggc cacccttttt gcctccatcc tcaaagacct ggtcttcaag tcatccgtca      60
gctgccaagt gttctgtaag aagatctact tcatctgggt gacgcggacc cagcgtcagt      120
ttgagtggct ggctgacatc atccgagagg tggaggagaa tgaccaccag gacctggtgt      180
ctgtgcacat ctacatcacc cagctggctg agaagttcga cctcaggacc actatgctgt      240
acatctgtga gcggcacttc cagaagggtc tgaaccggag tctattcaca ggctgctgt      300
ccatcaccca ctttggccgt cccctctttg agcccttctt caactccctg caggagggtcc      360
acccccacgt ccggaagatc ggagtgttta gctgtggcn      399

```

```

<210> 181
<211> 402
<212> DNA
<213> Homo sapiens

```

```

<400> 181
ggcacgaggc tacttcgctc gcaaggatta gtactgcaag gagtatgtgg cagctgtcct      60
ggagcatatc gagaacaaga acctcatgcc acctcttcta gtggtgcaga ccttggccca      120
catctccaca gccacactct gcgtcatcag ggactacctg gtccaaaaac tacagaaaca      180
gagccagcag attgcacagg atgagctgcg ggtgcggcgg taccgagagg agaccaccg      240
tatccgccag gagatccaag agctcaaggc cagtcctaag attttccaaa agaccaagtg      300
cagcatctgt aacagtgcct tggagatgcc ctcagtcac ttcctgtgtg gccactcctt      360
ccaccaacac tgctttgaga gttactcgga aagcgaagct ga      402

```

```

<210> 182
<211> 384
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(384)
<223> n = A,T,C or G

```

```

<400> 182
ggcacgagag caactcaggc ctgctggggt aactgcttac accattttcc ttcccctcct      60
cttccttgcc ttcgacactc ttaacctgga aaaagcacta atttgtcctc catatctgtg      120
gttttgtcat ttggaaaggt tgtagaaatc ctagagtatg tgacctttta agatgcactt      180
tttagaaaac tcaacatggt gctcttggtg taatagtttg ttcttttttag tgttcgggtat      240
tctcttggtg ggtcatgccc cagtttattt aacctacca tagatgttta tttcccttg      300
taaagttggt tagcatgtan nnnnnnnnnn nnnnnnggga aactcattct cnnnnnnnnn      360
nnnnnnnnnn nnnnntgccc cttg      384

```

```

<210> 183
<211> 384

```

<212> DNA

<213> Homo sapiens

<400> 183

ggcacgaggg	aaggtgaggg	ctgagaagga	ggcatgccag	cgggagaaaag	agctgcctgc	60
agcagtacat	cccttccatt	ttgttttaaat	tgggcttgga	gaatctattc	tgaaaacatt	120
gactctagac	ttgtagaaaa	gagccatttt	agtttcaact	caaagtataa	gcaaggtagt	180
ttggtgacat	tttgctttta	tgtgaaatag	tgcacagtat	gagttaatct	gagcaggtct	240
gaattgacca	aatgcttatc	tacgagggtc	ctagagctct	gctgaccctt	ggccgaaact	300
ctaaaatgta	cctattaaag	ataaatgctt	ctaccaaaag	aaaactctgt	gagttgtttc	360
agggcagaat	gtaccagcca	gtca				384

<210> 184

<211> 379

<212> DNA

<213> Homo sapiens

<400> 184

ggcacgagct	tcctccagcc	tccacagcct	tggctcagtg	tcctctgtgt	caagacccag	60
tgacttccag	gctcccagaa	accccaccct	aaccatgggc	caaccagaa	cacccactc	120
tccaccactg	gccaaagaac	atgccagcag	ctgccccca	tccatcacca	actccatggt	180
ggacataccc	attgtgctga	tcaacggctg	cccagaacca	gggtcttctc	cacccagcg	240
gacccagga	caccagaact	ccgttcaacc	tggagctgct	tctcccagca	acccctgtcc	300
agccaccagg	agcaacagcc	agaccctgtc	agatgcccc	tttaccacat	gccagaggg	360
tacgtcgtaa	accaatatt					379

<210> 185

<211> 368

<212> DNA

<213> Homo sapiens

<400> 185

ggcacgagac	ccggtccagg	tgccctacgt	cggcgcgagc	gcgcggcagg	tggagcacgt	60
gttgctgctg	ctgcgaggac	gccccgaaa	aacggtggat	ctgggctctg	gcgacggcag	120
gatcgtgctg	gcggcccaca	ggtgcggcct	ccgcccggcc	gtgggctacg	agctgaaccc	180
ctggctgggtg	gcgctggcgc	ggctgcacgc	ctggagggcc	ggctgtgccg	gcagcgtctg	240
ctatcgccgc	aaggatctct	ggaaggtaac	ctggggatcc	ctggccaccc	gctgacagcc	300
caaggtgcgg	ctgacacctg	cgagggctgg	gggccgggac	tcggaagctg	cgatgacccg	360
gtgcccac						368

<210> 186

<211> 375

<212> DNA

<213> Homo sapiens

<400> 186

ggcacgaggt	ctcacagagc	gagaaggtgt	caggagcagc	ccatttgtgt	ctctctctct	60
acctctgtga	agggcgcgaa	tgggcagagc	agaacttcta	gaagggaaga	tgagcaccca	120
ggatccctca	gatctgtgga	gcagatccga	tggagaggct	gagctgctcc	aggacttggg	180
gtggtatcac	ggcaacctca	cacgccatgc	tgctgaagct	cttctcctct	caaattggatg	240
tgacggcagc	taccttctga	gggacagcaa	tgagaccacc	gggctgtact	ctctctctgt	300
gagggccaaa	gattctgtta	aacactttca	tgttgaatat	actggatatt	catttaaatt	360
tggtctgta	gaatt					375

<210> 187

<211> 368

<212> DNA

<213> Homo sapiens

<400> 187

ggcacgaggg	cgtgcagagc	ctgtatggta	agcccctagg	gggctcaaag	gccggccagc	60
tcccaggaaa	gatgtgcact	gactttgaaa	cctgggactc	ctacagcccc	caagggaaggc	120
gccctgaaac	gcagggccct	aaatactgcc	actcttcctt	cgatgccatc	actgtagaca	180

ggcaacagca	actgtacatt	tttaaaggga	gccatttctg	ggaggcggca	gctgatggca	240
acgactcaga	gccccgtcca	ctgcaggaaa	gatgggtcgg	gctgcccccc	aacattgagg	300
ctgcggcagc	gtcattgaat	gatggagatt	tctacttctt	caaagggggg	cgatgctgga	360
ggatccgg						368

<210> 188
 <211> 436
 <212> DNA
 <213> Homo sapiens

<400> 188						
ggcagcagaa	ggggctgggg	tgggctcagg	caaggcctgg	ggccctggcc	tttttctctg	60
cagggggagg	caggggactg	tgagggggct	cagggaggcc	tccccacct	gccccctgac	120
cacaccact	ctgatgaggc	tcatggcctc	ctggcaggtc	gacggaggag	atcatcgccc	180
tcttcatttc	catcacgttt	gtgctggatg	ccgtcaaggg	cacggttaaa	atcttctgga	240
agtactacta	tgggcattac	ttggacgact	atcacacaaa	aaggacttca	tcccttgtca	300
gcctgtcagg	cctcggcgcc	agcctcaacg	ccagcctcca	cactgccctc	aatgccagct	360
tctctgccag	ccccacggag	ctgccctcgg	ccacacactc	aggccaggcg	accgccgtgc	420
tcagcctcct	catcat					436

<210> 189
 <211> 435
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(435)
 <223> n = A,T,C or G

<400> 189						
ggcagcagac	agaccctttc	ttcctaaagg	ctttgtggca	tcagacacat	aaaggggtata	60
tgtagtgtgg	agcactaacc	atggcagggt	aattttattcc	aggcacagag	tcataattct	120
ggaaacatct	agactcactg	cattaacaga	gcatttttgtt	tctaaagtag	acctcttatg	180
tcatccagat	ttcactcatt	ctgaccacag	ccaggaagct	gaggggtgaag	ccagaattag	240
ctgaaaccca	ccaagagctg	catagagcac	gttttagctag	agtaggagtt	tgcagtgtct	300
atatgggaaa	tgctgtctgt	atacttttag	gaattttctga	gtgcaattta	gaaacatcta	360
gcacacttga	aacactgcgt	atcattntcc	tcactcatga	atatagtcat	cagaattcat	420
aaatagttta	cctga					435

<210> 190
 <211> 437
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(437)
 <223> n = A,T,C or G

<400> 190						
ggcagcagat	taggaccctt	ccttggcaca	gggggtgaaa	agagcttggg	gaacgcttgg	60
cattatggag	ggctggaagg	ggctcaaccc	cgatttggag	agaagtttgg	gatggagtgg	120
gagagagatt	gagagagcga	gcaggaaaag	aggtcttggg	gcctgggact	gatggtggat	180
aaggcctgga	aagaagatga	cgaggaggag	gagagagggg	agtgggggtg	atgaggagca	240
ggctgacacc	tgggctgccc	tcaatcccca	aggccaggga	gggcggngct	ggccccctgg	300
aagaactggg	tctctgggct	ccctatgcac	tgcccaaact	ggctgagcca	ggagtggggc	360
aggaagtgag	agtcaaggcc	cagcaaaagg	agggggagga	gctgccaatt	ataaccttgt	420
gganggaccg	gttttgn					437

<210> 191
 <211> 434
 <212> DNA

<213> Homo sapiens

<400> 191

ggcaccgagaa	gaaactgtga	agggaaagaa	aggtttatac	tgagaaatgg	aagagataat	60
tttagaaaact	tgtgaaaaat	ggcttaaatct	aaatgagtgt	taggggagat	acagttgtga	120
tgatagggttg	agctcacatg	gtggagagcc	acagttgcgg	gtgcttgcac	tgataatgtg	180
agggcatgga	gacagacaat	aggttgaatg	ctcttttttt	acaaaaggaa	gtagaaaggg	240
agggggatgt	aaatttgata	aatagggttg	tgaaaactta	tattttcttg	taaagagaga	300
gaactgagca	tgttgtaggt	ataaggtaaa	aaggcgtgaa	gaggaatatt	tcgttgataa	360
tgaaagttag	cagctaggga	agaaaactcc	cagaggaaga	gggaggcaag	gaaatcaaga	420
acacacttaa	agtg					434

<210> 192

<211> 323

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(323)

<223> n = A,T,C or G

<400> 192

gggtctctcg	ccccctctc	tctcttttgt	gtgtctctct	ctctgtcccg	tgtgtgnnnn	60
nnnnnnnnnt	ctctctatat	ctcgcgcgcg	cgcactcccg	tgtgtgtgtg	tgaccccgcc	120
ccctcatgcg	ctctctcatt	tgtggagaga	gagaccgcta	tctatctctc	tctccccgcg	180
ccatacacaca	tctccctctc	tgtgaaagag	acgtgtgtgt	gtctccacac	cccttgggcg	240
cgcgcgcgcc	acccccctctc	ctgggggggg	tgtcctctct	gtatatatat	atgtgcacac	300
acgcgcgcgc	gctctgtggt	gtt				323

<210> 193

<211> 412

<212> DNA

<213> Homo sapiens

<400> 193

ggcaccgagaa	ggggccgtga	cagccgttgc	catctgctgc	cggagccggc	acctggcgca	60
ggcctcccgag	gagctccagt	gacagcccca	tcccaggatg	gggtgtctggg	gaggggtcaag	120
ggctgggggct	gagctttaaa	atggttccga	cttgtccctc	tctcagccct	ccatggcctg	180
gcacgagggg	atggggatgc	ttccgccttt	cgggggctgc	tgccctggcc	cttgagtggg	240
gcagcctcct	tgccctggaac	tcaactcactc	tgggtgcctc	ctccccaggt	ggaggtgccca	300
ggaagctccc	tccctcactg	tggggcattt	caccattcaa	acaggtcgag	ctgtgctcgg	360
gtgctgccag	ctgctcccaa	tgtgccgatg	tccgtgggca	gaatgacttt	ta	412

<210> 194

<211> 405

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(405)

<223> n = A,T,C or G

<400> 194

cgttgctgtc	ggtcagcaat	gaaataaata	tctttagtaa	tgttcnnnnn	nnnnnnnnnn	60
ngaaccctc	gggggcccctt	ttttcccgaa	acccccactg	gaaaaaaacc	cttgggggggt	120
tggcaaaacc	ccccataaaa	agggggggaa	aaaaaggctt	tttttggaag	aatggggggg	180
tctttgcttt	ttttggaccc	ctttaagcg	gggaaaacca	ggttaacccc	ccccaggggc	240
nnnnnnnnnn	gtttcagggc	cnnnnnnnnn	nnnnnnnnnt	tttttccttn	tctcccttct	300
gtctcgccct	gctgcgctgc	cgttttctcg	ttccactccc	cccgtttttg	tactcccccc	360
gtgccggtga	gcgtccaccc	tattctttcg	cgccggtgca	cccc		405

<210> 195
 <211> 400
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(400)
 <223> n = A,T,C or G

<400> 195
 ggcacgagat taggaccctt ccttggcaca ggggtgagaa agagcttggg gaacgcttgg 60
 cattatggag ggctggaagg ggctcaaccc cgatttggag agaagtttgg gatggagtgg 120
 gcgagagatt gagagagcga gcaggaaaag aggtcttggg gcctgggact gatggtggat 180
 aaggccttga aagaagatga cgaggaggag gagagagggg agtgggggtgg atgaggagca 240
 ngctgacacc tgggctgccc tcaatcccca aggccaggga gggcgngct ggccctggg 300
 aagaactggg tctctgggct ccctaggcac tgcccaaact ggctgagcca ggagtggggc 360
 aagaaatgag agttcaggcc caacacaagg agggggaggg 400

<210> 196
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 196
 ggcacgagat taggaccctt ccttggctca ggggtgagaa agagcttggg gaacgcttgg 60
 cattatggag ggctggaagg ggctcaaccc cgatttggag agaagtttgg gatggagtgg 120
 gcgagagatt gatagagcga gcaggaaaag aggtcttggg gcctgggact gatggtggat 180
 aaggccttga aagaagatac taggaggagg agagagggaa gtgggggtgga tgaggagcag 240
 gctgacacct gggctgccct caatcccca ggccaggag ggcggggctg gcccctggga 300
 agaactgggt ctctgggctc cctaggcact gcccaaactg gctgaaccag gagtggggca 360
 agaagtgaga gtcaaggccc aacaaaagga gggggaggag ct 402

<210> 197
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 197
 ggcacgagct ctcagcggcc ggtttctgcg tccgctgccg caggttccac cgcgctccag 60
 gtattttttt ttctgaagga aagctgcttc ctcatatgtt tcaagaatgg ctctccctat 120
 cattgtaaaa tggggtggac aggagtattc agtgaccaca ctttcagaag atgatactgt 180
 gctcgatctc aaacagtttc tcaagaccct tacaggagtt cttccagaac gccaaaagtt 240
 acttggaactc aaagttaaag gcaaaccctgc agaaaatgat gttaagcttg gagctctcaa 300
 actgaaacca aatactaaaa tcatgatgat gggaactcgt gaggagagct tggaagatgt 360
 cttaggtcca cccctgaca atgatgatgt tgtaatgac t 401

<210> 198
 <211> 397
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(397)
 <223> n = A,T,C or G

<400> 198
 tgcataattag acattcttaa cagggcggca gtctagtgtt gaaagtttta tttttccatt 60
 tttcttttaa gcaaattttt tttaaaaaat tctgattnnn nnnnnnnnnn nnnnnnnnnn 120
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn tctgatttaa 240
 ttcttttatt tatcataagg ggtttaattc ctgaagtaaa ggtttgcacc tattaaactt 300

aaaactgccca	aatgattttt	gttcttttat	gtgcgcgata	gaaatacaaa	gaatggagt	360
gccacctcct	ccctttcaag	ctagggcagc	agggacg			397

<210> 199

<211> 398

<212> DNA

<213> Homo sapiens

<400> 199

ggcacgagaa	gaaaggttta	tactgagaaa	tggaagagat	aatttttagaa	acttgtgaaa	60
aatggcctaa	tctaaatgag	tgtagggga	gatacagttg	tgatgatagg	ttgagctcac	120
atggtggaga	gccacagttg	cgggtgcttg	caactgataat	gtgagggcat	ggagacagac	180
aataggttga	atgctctttt	tttacaaaag	gaagtagaaa	gggaggggga	tgtaaatttg	240
ataaatagggt	tggtgaaaac	ttatatatttc	ttgtaaagag	agagaactga	gcatgttgta	300
ggtataagggt	aaaaaggcgt	gaagaggaat	atttcgttga	taatgaaagg	gagcaactta	360
gggaaaaaaa	cttcccaagg	aggaggggag	cagggaaa			398

<210> 200

<211> 394

<212> DNA

<213> Homo sapiens

<400> 200

ggcacgagca	gaaggcagcg	gtctaggcga	ggacgcccgg	ctggaccagg	agaccgcca	60
gtggctgcgc	tgggacaaga	attccttaac	tttggaggca	gtgaaacgac	taatagcaga	120
aggtataaaa	gaagaactac	gaaaatgttt	tggggcccgga	atggagtttg	ggacagctgg	180
cctccgagct	gctatgggac	ctggaatttc	tcgtatgaat	gacttgacca	tcattccagac	240
tacacagggga	ttttgcagat	acctggaaaa	acaattcagt	gacttaaagc	agaaaggcat	300
cgtgatcagt	tttgacgccc	gagctcatcc	atccagtgga	ggtagcagca	gaaggtttgc	360
ccgacttgct	gcaaccacat	ttatcagtca	gggg			394

<210> 201

<211> 391

<212> DNA

<213> Homo sapiens

<400> 201

ggcacgagca	ggcgtgtctg	ggtaaccatg	tggtcctctg	tggtcctccc	tgctgtctcc	60
caaagcacag	ggctcagctc	cagagggaga	cgggctgggc	tgctcagtggt	cccaggtgca	120
tcccactttc	cagcagcact	tggtgccagc	agaggctgca	ggtgtggcag	gagggggccc	180
agccgtgagg	gcaccagggt	caggcccggc	atctcagggt	ggagagccag	ggctgtcctg	240
aacctccaga	gggggtgagc	tgggaaacttg	tgtgaagggg	ctttttccaa	aaggaaaaacg	300
ggagcttact	ggctcacggc	tgatgcccc	gacagcctcg	aggatctgca	ggtccccaga	360
caccaagcct	gggtgctctc	cagcagacgg	c			391

<210> 202

<211> 392

<212> DNA

<213> Homo sapiens

<400> 202

ggcacgagat	tctcagtaca	ctaaacactt	gttaagagtg	ttgttaagag	ccagagtggg	60
tatcatgtgg	gacacagacc	ctttcttcct	aaaggctttg	tggtatcaga	cacataaagg	120
gtatatgtag	tgtaggagc	taaccatggc	agggtaat	attccaggca	cagagtcata	180
attctggaaa	catctagact	cactgcatta	acagagcatt	ttgtttctaa	agtagacctc	240
ttatgtcatc	cagatttcac	tcattctgac	cacagccagg	aagctgaggg	tgaagccaga	300
attagctgaa	accaccaag	agctgcatag	agcacgttta	gctagagtag	gagtttgcag	360
tgctcatatg	ggaaatgctg	ctgctatact	tt			392

<210> 203

<211> 392

<212> DNA

<213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(392)
 <223> n = A,T,C or G

<400> 203
 ggcacgagga ggagcccgcc ccggaggctg aggctctggc cgcagcccgg gagcggagca 60
 gccgcttctt gagcggcctg gagctgggtga agcaggggtgc cgaggcgcgcg gtgttccgtg 120
 gccgcttcca gggccgcgcg gcggtgatca agcaccgctt ccccaagggc taccggcacc 180
 cggcgctgga ggcgcggctt ggcagacggc ggacgggtgca ggaggcccg ggcctcctcc 240
 gctgtcgccg cgctggaata tctgccccag ttgtcttttt tgtggactat gcttccaaact 300
 gcttatatat ggaagaaatt gaaggctcag tgactgttcg agattatatt cagtccacta 360
 tggagactga aaaaactccc cagggtctct cn 392

<210> 204
 <211> 386
 <212> DNA
 <213> Homo sapiens

<400> 204
 ggcacgagaa gccttaaaccc gggaaatttc catgctatct agagggttttt gatgtcatct 60
 taagaaacac acttaagagc atcagattta ctgattgcat tttatgcttt aagtacgaaa 120
 gggtttgtgc caatattcac tacgtattat gcagtattta tatcttttgt atgtaaaact 180
 ttaactgatt tctgtcattc atcaatgagt agaagtaaat acattatagt tgatttttgc 240
 aaatcttaat ttaaaagcct cattttccta gaaatctaatt tattcagtta ttcattgacaa 300
 tatttttttta aaagtaagaa attctgagtt gtcttcttgg agctgtaggt cttgaagcag 360
 caacgtcttt caggggttgg agacag 386

<210> 205
 <211> 295
 <212> DNA
 <213> Homo sapiens

<400> 205
 gcgctctctt cacacacaaa agatatatat atagaaaggg agtgtggata tccccctaa 60
 atatgtgagc gtgtctctct cgaccgtctc cccagagaaa aatatctcta gagagagcac 120
 aagtgtgttc tctgtgtctt gtgtgtgaga aaaaataagt gcccgcgcac acatagattt 180
 ttatatcgct cccccccgcg cctttatata tgttttttgg gtgtatatat attttataca 240
 aaaacatggt tcttttttgag gcccttaca acaaaaattt tgttcttttt gaacc 295

<210> 206
 <211> 383
 <212> DNA
 <213> Homo sapiens

<400> 206
 ggcacgaggt taccatcag cccttgcaag tccccactc aggcctctgg aaggtccagg 60
 gatgggctct gatgagaggg taaaagatgc tcagggaaac acaggcctca gctgcctaga 120
 ggaccctccc cctgccttgc agtgggctcg ggtagagcag tatcaggagc tagggttgtc 180
 tgctgcccac actcctgctt tttgggatat ctaactgcta aggagggagt tgacatcccc 240
 ctctctggctc atgtgtctga caccaacaac atgggtctctg tccctctctc tttgactctc 300
 cctttgtcct ccccatagag ctgggggtgg gtggatccct atacctgggg caggcagccc 360
 caaagtgggg gaggggggatg gca 383

<210> 207
 <211> 385
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(385)
 <223> n = A,T,C or G

<400> 207
 ggcacgagct tcaggataag aagctcatgg ccatgttctt agagtataac aaagccatcc 60
 ggaactacac ccgcttcgat gactgggtacc tgtgggttca gatgtacaag gggactgtgt 120
 ccatgccagt cttccagtc tggaggcct actggcctgg tcttcagagc ctcatggag 180
 acattgacaa tgccatgagg accttctctca actactacac tgtatggaag cagtttgggg 240
 ggctcccggg attctacaac attcctcagg gatacacagt ggagaagcga gagggtacc 300
 cacttcggcc agaacttatt gaaagcgcaa tgtacctcta ccgtgccacg gnggatccca 360
 ccctcctaga actcgaaga gatgg 385

<210> 208
 <211> 374
 <212> DNA
 <213> Homo sapiens

<400> 208
 ggcacgagcc tcagctgcct agaggaccct cccctgcct tgcagtgggc tcgggtagag 60
 cagtatcagg agctaggggt gtctgctgcc cacactcctg ctttttggga tatctaactg 120
 ctaaggaggg agttgacatc ccccttctgg ctcatgtgtc tgacaccaac aacatggtct 180
 ctgtccctct ctctttgact ctccctttgt cctcccata gagctggggg ggggtggatc 240
 cctatacctg gggcaggcag ccccaaagtg ggggagggg atggcagaga ctgtaaaggc 300
 gccactggac tctggcaagg cctttattac ctttactccc ctccctctcc catcaccagc 360
 ctcaaggcct gagg 374

<210> 209
 <211> 425
 <212> DNA
 <213> Homo sapiens

<400> 209
 ggcacgagcc caagtgcctt ctgcagaggt tgtcgttga aaactgtcac cttacagaag 60
 ccaattgcaa ggaccttgct gctgtgttg ttgtcagccg ggagctgaca cacctgtgct 120
 tggccaagaa ccccatggg aatacagggg tgaagtttct gtgtgagggc ttgaggtacc 180
 ccgagtgtaa actgcagacc ttggtgcctt ggaactgcga cataactagc gatggctgct 240
 gcgatctcac aaagcttctc caagaaaaat caagcctgtt gtgtttggat ctggggctga 300
 atcacatagg agttaaggga atgaagttcc tgtgtgagggc tttgaggaaa ccactgtgca 360
 acttgagatg tctgtggttg tggggatgtt ccatccctcc gttcagttgt gaagacctct 420
 gctct 425

<210> 210
 <211> 396
 <212> DNA
 <213> Homo sapiens

<400> 210
 ggcacgagga gcaaggaagt aatattgtca tatttgcagt tgagaatgat ccctgagtct 60
 cggtttttctt atctatgaaa tgaggctaag aataataaaa tagagaatta aatgagataa 120
 tgctgtataa cagtgcctgg catatagctt attattcatc cagctaagag gcccttccat 180
 atgtgaagct ttgtctctgt aggtctgtat tacaatcaca ttcagttata gctaattatt 240
 tacttatgta gctatctctg aaacttagaa atgaaatcat cgaggaaaaa ggccatttct 300
 tgatcctgtc tgtgttccct gttcccagca taaagcctaa cacgtattag gctaattgtc 360
 ccgagcaaag aaagcatcaa agtggcgggt cgggcc 396

<210> 211
 <211> 267
 <212> DNA
 <213> Homo sapiens

<400> 211
 tctctagaga cacacagaga gggtagcggt ctctctcaca cgcaccccag agtcaggcgc 60
 gcacgctctc tctctctctc tctatccctc agaaagatct tcctttttcc ctctccctgt 120
 gatgtagtga gagtttgatg catatttgct cgtgtccgcc cccacagacc ctctacctct 180
 ctgtgctggc cctatcttgt gtgtatgttt ccctctctct ctccgcgcgc cacacgatgt 240
 actttcttta tatgtagtgc cagttcc 267

<210> 212
 <211> 396
 <212> DNA
 <213> Homo sapiens

<400> 212
 ggcacgagcc aggaggaccc tcgcttcctc tccgccatgc ttgccacctc ttgcttctga 60
 gagtccatct cagttcgcag ttctgtgact tgcattgacc ttgctccaat caagctacaa 120
 ctcaagcagt cacggggaga aggattgtag atggggcagt gactcacagg gtcaggcact 180
 cgggggagcc tgagtcagga ggtcagtggg ccctggaagg gagggggcaa gcctgggtgg 240
 gtaaggttct gggccccagg caagaaggca gagtttctcc gcaggggtgt gtgcaagagc 300
 tagctgcgca gaaggtctcc gctggctctc caagccgggc ttgtgaaata ggaacgccaa 360
 catcctctc caccaggcagt ggcaggcacc tcctcc 396

<210> 213
 <211> 284
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(284)
 <223> n = A,T,C or G

<400> 213
 tgggctgtct cgcacctcct ccctctctct ttgtactcac agtgaaaaat tatagtgttc 60
 gcgtgcgggg cgcgctcttt actttttttt ctctctcaca catatttata tatatagaga 120
 gagcctccga gcgctctgcc cccctcctct ctctctctct tcacgtgtgt gcatcaccca 180
 ctcnnnnnnn nnnnctcttc cagagatacg ggggcttgtt tcctccgctc tctctcacac 240
 gtctgtgcag cagaggacta tttttttctt tccccgcgt ctcn 284

<210> 214
 <211> 440
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(440)
 <223> n = A,T,C or G

<400> 214
 ggcacgaggg attgcagtca gcactttctg aatgttttca cacagtatgc aaagcttaca 60
 tcataccaag gagtggagag ttgaagtttc ctcccagtga ctccagtgc agaccacacc 120
 tagaaaagcgt ttctcttcct gattatttca aaaagatgta aaagagctgg ggagagtatg 180
 ggaagaaaca atacaggatt gcctttaatt aattaagaat tgcctcctga taaaaggaaa 240
 aagaaattaa tgctggagta tggaggggtg ataaccttaa agattataaa tatttgttgt 300
 ctataaatat ttataaatta taaacacaat ataattaaaa ttagaacatc aggaaaagaa 360
 ttaaaatcct cagggttgcaa aaccaaagt ttaaccaaaa caaatactca tgagattcaa 420
 ctttgtttcac ctatagaaan 440

<210> 215
 <211> 439
 <212> DNA
 <213> Homo sapiens

<400> 215
 ggcacgagtgc cacaggggac acttacggac acagaaatgc acaggggagg ccgagcataa 60
 ccaggggtga ggggcaggca gcagttgtag ttactgccgc ggggcactgc tatgtgcagg 120
 gacagccagc acccagccca tcaccactcc ctgggctggc tggcaggtat ggcaccctgg 180
 gagcccgga tataccagg gcaccctac ggctgccgc agtctcatgc ccaggtgggt 240
 gctctgggct ggagcgagg ccaggttttg ggccgaggct tccccaggca atcctgtgag 300
 ctcccttcta gcctctgacc cagtctggtc tggcttgcac ggatgtaggg cttgggggtg 360

gaagttcagg	tcctggcttt	gcctttgcct	gatgtggatg	agcagctcac	atgctcaggg	420
ccacctgaga	ctgtcactg					439

<210> 216
 <211> 392
 <212> DNA
 <213> Homo sapiens

<400> 216						
ggcacgagga	gacagagaag	tttggccagg	gggtccacca	tactgctggt	caggttggga	60
aggaggcaga	gaagtttggc	caggtgggga	aggaggaaga	cagagtgggc	caaggcctcc	120
atcatggcgt	tagtcaggct	ggaagggagg	cggggcagtt	tggccacgac	attcaccaca	180
cagcagggca	ggctgggaaa	gaggagagaca	tagcagttca	tgggtgtccaa	cctggggtcc	240
acgagggcgg	gaaggaggca	gggcagtttg	gccagggagt	tcaccatacc	cttgaacagg	300
ccgggaagga	agcagacaaa	gcggtccaag	ggttccacac	tgggggtccac	caggctggga	360
aggaagcaga	gaaacttggc	ccaggggtca	ac			392

<210> 217
 <211> 394
 <212> DNA
 <213> Homo sapiens

<400> 217						
ggcacgagcc	catctggggc	agcaccacgt	ggatctctcc	ctcgtcacct	tcaactgggt	60
cctcgtggtc	tttgcgga	gtctcattag	caacatcctc	cttcgggtct	gggatgcctt	120
cctgtacgag	gggacgaagg	tgggtgttcg	ctatgccttg	gccattttca	agtacaacga	180
gaaggagatc	ttgaggctac	agaatggcct	ggaaatctac	cagtacctgc	gcttcttcac	240
caagaccatc	tccaacagcc	ggaagctgat	gaacatcgcc	ttcaatgaca	tgaacccctt	300
ccgcatgaaa	cagctgcggc	agctgcgcac	ggtccaccgg	gagcggctgg	aggctgagct	360
gcgggagctg	gagcagctta	aggcagagta	cctg			394

<210> 218
 <211> 432
 <212> DNA
 <213> Homo sapiens

<400> 218						
acaccacatt	gtttgaggac	accatcgatt	cgaattcggc	acgagcctag	ccagcccctg	60
acgtgcctta	caggagttct	tccagacacg	ccaaaagtga	cttggactca	aagttaaagg	120
caaacctgca	gaaaatgatg	ttaagcttgt	agctctcaaa	ctgaaaccac	ataactaatat	180
catgaggatg	gcatctcgag	aggagagctt	ggaagatgtc	ttaggtccac	cccctgacaa	240
tgatgatgtt	gttaatgact	ttgatattga	agatgaagta	gttgaagtag	aaaatagggg	300
agaaaaccta	ctgaaaat	ctcgcagagc	gaaagagtac	aaagtggaaa	ttttgaatcc	360
tcccaggga	gggaaaaagc	ttttggtgct	agatgttgat	tatacattat	ttgaccacag	420
gtcttgtgca	ag					432

<210> 219
 <211> 395
 <212> DNA
 <213> Homo sapiens

<400> 219						
ggcacgagcc	ctttactcct	ctacccaaga	tcttgcttgt	ttctttctaa	gttgcctctc	60
tatctagctt	gcaggatttg	agttgaggaa	aacacagact	tccatgagtt	tgggaactac	120
gagagaaaag	acagacagag	tcaaactctac	agcatatctc	tcacctcagg	aactggaaga	180
tgtattttat	caatatgatg	taaagtctga	aatatacagc	tttggaatcg	tcctctggga	240
aatcgccact	ggagatatcc	cgtttcaagg	ctgtaattct	gagaagatcc	gcaagctggg	300
ggctgtgaag	cggcagcagg	agccactggg	tgaagactgc	ccttcagagc	tgcgggagat	360
cattgatgag	tgccgggccc	atgatccctc	tgtgc			395

<210> 220
 <211> 487
 <212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(487)

<223> n = A,T,C or G

<400> 220

tgctctttttg	atgatgccat	cgattcgaat	tcggcagcag	cagctagctc	agttcaaggt	60
ggaaatggct	taacgagagg	aacggcaaca	gcaggtggct	gaggactacg	agctcagact	120
ggcccgggag	caagcgcgag	tgtgcgaact	gcagagtggg	aaccagcagc	tggaggagca	180
gcgggtggag	ctggtggaaa	gactgcaggc	catgctgcag	gccactggg	atgaggccaa	240
ccagctgctc	agcaccactc	tcccgcgcgc	caaccctcca	gtcctcctcg	ctggaccctc	300
cagccccggg	cctcaggagc	ccgagaagga	ggagaggagg	gtctggacta	tgccctccat	360
ggccgtggcc	ctgaagcctg	tattgcagca	gagccgggaa	gcaagggacg	agctacctgg	420
agcgctcctc	ggtttttgca	gntcctcctc	agatcttagc	ctcctgggtg	gccccctctt	480
tcagagc						487

<210> 221

<211> 365

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(365)

<223> n = A,T,C or G

<400> 221

ggatgccagt	ggtgaggctg	taagcgaaac	tcttcagttt	aaagctcaag	atctcttaag	60
ggcagtccca	agatccagag	cagagatgta	tgatgacgtc	cacagcgatg	gcagatactc	120
cctcagtgga	tctgtagctc	actctagaga	tgccggaaga	gaaggcctga	gaagtgacgt	180
atttccaggg	ccttccttca	gatcaagcaa	cccttccatc	agtgatgaca	gtacttttcg	240
caaagaatgt	ggccgggagc	tggaattttc	tcactctgat	tctcgggacc	aggtcattgg	300
ccaccggaaa	ttggggcatt	tccgtttctc	ggactggaaa	tttgcgctcc	gtggttcttg	360
ggaan						365

<210> 222

<211> 376

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(376)

<223> n = A,T,C or G

<400> 222

ggcagcagga	gatttcccgg	cgggtcccgg	cctctgcgtg	cacgcgcctg	cgtgctcgcg	60
ctcgcggttc	tggcgctgct	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	120
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	180
nnnnnnnnnn	nnntgatggc	agcagtggcc	tgccctgaagc	agccactgcc	aagaatctag	240
ctgcagtnnn	nnnnnnnnnn	nnnnnnnnca	tgctccacac	agccaccgga	agccaagaac	300
gcaccctcct	gggtacagct	gcaagccgcc	agccgaggct	gcggaccggg	gcctccctgg	360
tgctctgggg	gttggg					376

<210> 223

<211> 399

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(399)

<223> n = A,T,C or G

<400> 223

ggcacgaggg	gtgacagagc	ggctggcgca	tgctcagtag	agcagcctac	ggcaagcagc	60
ctccctcagg	gaacatcaca	ggaagcagct	gcaggacctg	agtggacagc	accagcagga	120
gctggccagt	cagctagctc	agttcaaggt	ggaaatggca	gaacgagagg	aacggcaaca	180
gcaggtggct	gaggactacg	agctcagact	ggcccgggag	caagcgcgag	tgtgcgaact	240
gcagagtggg	aaccagcagc	tggaggagca	gcgggtggag	ctggtggaaa	gactgcaggc	300
catgctgcag	gcccactggg	atgaggccaa	ccagctgctc	agcaccactc	ttccgcccgc	360
caaaccttca	gctttcttctg	cttgaccctc	cagccccgn			399

<210> 224

<211> 402

<212> DNA

<213> Homo sapiens

<400> 224

ggcacgaggg	cagttcagta	tcgatggaca	gatcttccta	ctctttgact	cagagaagag	60
aatgtgggca	acggttcata	ctggagccag	aaagatgaaa	gaaaagtggg	agaatgacaa	120
ggatgtggcc	atgtccttcc	attacatctc	aatgggagac	tgcataggat	ggcttgagga	180
cttcttgatg	ggcatggaca	gcacctgga	gccaaagtga	ggagcaccac	tcgccatgtc	240
ctcaggcaca	acccaactca	gggccacagc	caccaccctc	atcctttgct	gcctcctcat	300
catcctcccc	tgcttcatcc	tccctggcat	ctgaggagaa	tccttttagag	tgacagggtta	360
aagatgatac	caaaaagccc	ctgtgagcac	ggtcttgatc	ag		402

<210> 225

<211> 270

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(270)

<223> n = A,T,C or G

<400> 225

ctctctttct	ttctccctcc	ccccccgggc	gcgctcattt	atctcgtctc	ttatgtctct	60
ctctctgtgt	ctgtgacaga	cacactcttt	ttcatatagc	gcgctccctt	ttctttgctc	120
tcgggggggg	tctctctgta	cgcgtgtgtt	ctctctccag	tgagtgtgca	cgcctagggtg	180
agagagagtn	nnnnnnnnnn	nnnnntgtgt	gtgaatttta	tatatattcta	tatctctcac	240
tctctgggtg	tcacactctc	cgtgtgtggg				270

<210> 226

<211> 404

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(404)

<223> n = A,T,C or G

<400> 226

ggcacgagaa	ccctcccagg	ctaagcccca	atttggggct	cgcctgccct	gcatcagggg	60
gacatgtcag	ctgaggagta	attgaccaga	tttctgcttt	agaaatatgg	cagtggaggc	120
aggagatggc	atctgaggcc	caggctgggg	agaagggtgc	tgggatgaga	acctggagtt	180
cagaccaggg	aagggatgag	agcctaagaa	gaggagctct	cacctgaga	caggctgggtg	240
caggagtctg	ctcgatccag	gcctgggtcc	ctggttccct	ctgagcttgg	gaggactatg	300
tgagacagaa	caggaccagg	ggcctgcatt	cccccttgta	ttattcatct	tcnnnnnnnn	360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnn		404

<210> 227

<211> 389
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(389)
 <223> n = A,T,C or G

<400> 227
 ggcacgagaa gtcactcaac ctctctgagc cttttcttca cctataaagt ggggatagta 60
 actacctacc ttatggaagc atatgaggat tgtgtgaaat catccatgta gcccttccac 120
 cgccacgtgg agtttgatg ggagcagttt ctaaattgaa gtcattctga tcaggtgggc 180
 tgccaacctc tctgagcctc agtttgctct tctagggaat ggggacaatg caatgggaat 240
 ctgaggattg tgtgaaattg tgcaaatgca tgaatgtggg ctgggatagt aaaagggagg 300
 gccccggagc agccacactg gggtcctatc tagtggacgc gcccggtgcc caccattgc 360
 tgtgatgcc gacagccact gcaagcatn 389

<210> 228
 <211> 384
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(384)
 <223> n = A,T,C or G

<400> 228
 ggcacgagct gccacctcta gaaagctgct ttcttctatc accgcttgcc cttgaattat 60
 tccctgaatg aagccaagaa ccctcccagg ctaagcccca atttgggggt cgccctgccct 120
 gcattcaggga gacatgtcag ctgaggagta attgaccaga tttctgcttt agaaatatgg 180
 cagtggaggc agagatggc atctgaggcc caggctgggg agaaggggtg tgggatgaga 240
 acctggagtt cagaccaggg aagggatgag agcctaagaa gaggagctct caccctgaga 300
 caggctgggtg caggagtctg ctcgatccag gcctgggtcc ctggttccct ctgagcttgg 360
 gaggactatg tgagacagaa cagn 384

<210> 229
 <211> 292
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(292)
 <223> n = A,T,C or G

<400> 229
 ggtgtctctc tctcgggggg gccccccctc tctctatctt tttttgcgcg cactactcact 60
 ctctctctct tttttccccc gcgcgcgcgc acgcgctttt tttttctttt ttctnnnnnn 120
 nnnnnactct ctctcttttc tcttttgtgt gggggtctcc ggcgcgcttc tctctctctc 180
 tctcaccac agacactctc tctgtgtgcg cactctctc tctcgggggg ccggatctct 240
 cteccccctc tctatctctg ttattttggg ggtccctcc gcgctctct ca 292

<210> 230
 <211> 400
 <212> DNA
 <213> Homo sapiens

<400> 230
 ggcacgaggt gggacagaag tagaagagg tgaatggccc tggcaggcta gcctgcagtg 60
 ggatgggagt catcgctgtg gagcaacctt aattaatgcc acatggcttg tgagtgtctg 120
 tcaactgtttt acaacatata agaacctgc cagatggact gcttcctttg gagtaacaat 180

aaaaccttcg	aaaatgaaac	ggggtctccg	gagaataatt	gtccatgaaa	aatacaaaaca	240
cccatcacat	gactatgata	tttctcttgc	agagctttct	agccctgttc	cctacacaaa	300
tgcagtacat	agagtttgtc	tccctgatgc	atcctgtgag	tttcaaccag	gtgatgtgat	360
gtttgtgaca	ggatttgag	cactgaaaaa	tgatggttac			400

<210> 231
 <211> 332
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(332)
 <223> n = A,T,C or G

<400> 231						
tatatagaca	ccccgccttt	tttctctctc	tctctatata	cacacccgtct	ctctcccccg	60
tgtgtctctc	ccctctcttt	tgctcatact	tatatacata	tacacacttg	tgtgggggac	120
tctctctagc	gctccctctc	ttttgtgtgg	gcgctctcac	acacacacac	nnnnnnnnnn	180
nnagagactc	ctttctctgt	ggagaatatg	tgtgcgcacc	atctctctct	ctcttatttt	240
tccctcgcgc	gcgcgctctg	tgagagagac	tctctgttct	cacacatatg	atatatatat	300
ccctccccctc	tctcacactc	gtgccccgcg	cn			332

<210> 232
 <211> 407
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(407)
 <223> n = A,T,C or G

<400> 232						
ggcacgagaa	ctccggctac	gtttgtctgc	ccaacaaata	gaccaggggt	ccctaagtgt	60
cgcttcctcc	aagaagccct	ccctgatgag	ttgagccact	ttagtttgtg	ctcaggctca	120
ccctgcacgt	cttggttgct	ctcatcactg	taatgatcta	aaacacacgt	ctgctcatga	180
gacccgcac	ccacccccga	tgctggggcc	gctcttggat	tttcatgcct	gctgccagca	240
cccaggggga	gctccggaaa	tgtctgctgg	gggctcggaa	taccacacct	tctggtaatg	300
cagcccagcg	ggccccagcc	tcgttttcca	gccctcactc	anaatggagt	cgctctgggt	360
cgaacgcctc	tgancagtgt	gtacctacgt	gtcaggccca	tccttcc		407

<210> 233
 <211> 406
 <212> DNA
 <213> Homo sapiens

<400> 233						
ggcacgagga	aagaccacg	tgctgcctca	tgtggccgac	atcctcagca	agtcttgccc	60
ggcaccacag	tgagcctctg	gtgggggtgg	gtagtcacca	ctcggtctctg	gaggatgagg	120
cctggggccat	aatccagttg	cagggacgga	tgatctccat	ctcgaaggct	ccagaggtaa	180
ctgcgttgtc	ccatcctcca	ggcatcccct	gcggcgctgg	ccaagtgcgt	gctggccgag	240
gtcccgaagc	aggtggtgga	gtactacagc	cacagaggcc	tgcccccgag	aagcctgggt	300
gtccctgccg	gagaggccag	cccaggctgc	acaccgtgaa	aatgtggagg	gcgtaaaggg	360
ggggcccaga	aagaaagtgt	cccacacaa	ctctgtttgc	acatgg		406

<210> 234
 <211> 380
 <212> DNA
 <213> Homo sapiens

<400> 234						
ggcacgagga	gggtgaatgg	ccctggcagg	ctagcctgca	gtgggatggg	agtcacgcgt	60

gtggagcaac	cttaattaat	gccacatggc	ttgtgagtgc	tgctcactgt	tttacaacat	120
ataagaaccc	tgccagatgg	actgcttcct	ttggagtaac	aataaaacct	tcgaaaatga	180
aacgggggtct	cgggagaata	attgtccatg	aaaaatacaa	acacccatca	catgactatg	240
atatttctct	tgccagagctt	tctagccctg	ttccctacac	aaatgcagta	catagagttt	300
gtctccctga	tgcatcctat	gagtttcaac	caggtgatgt	gatgtttgtg	acaggatttg	360
gagcactgaa	aaatgatggt					380

<210> 235

<211> 410

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(410)

<223> n = A,T,C or G

<400> 235

ggcagcagct	gagcaggact	tagaggaact	ccggctacgt	ttgctgtccc	aacaaataga	60
ccaggggttcc	ctaagtgtcg	cttcctccaa	gaagccctcc	ctgatgagtt	gagccacttt	120
agtttgtgct	caggctcacc	ctgcacgtct	tggttgctct	catcactgta	atgatctaaa	180
acacacgtct	gctcatgaga	cccgcacccc	acccccgatg	ctggggccgc	tcttggattt	240
tcatgcctgc	tgccagcacc	cagggggagc	tccggaaatg	tctgctgggg	gctcgggaata	300
cccacctttc	tggtaatgca	gccagcggg	tcccagcctc	gttntccagc	cctcactcan	360
aatggagtgc	ctctggttcg	aacgcctctg	acaagtgtgt	acctacgtgt		410

<210> 236

<211> 394

<212> DNA

<213> Homo sapiens

<400> 236

ggcagcagac	tccggctacg	tttgctgtcc	caacaaatag	accagggttc	cctaagtgtc	60
gcttcctcca	agaagccctc	cctgatgagt	tgagccactt	tagtttgtgc	tcagggtcac	120
cctgcacgtc	ttggttgctc	tcatcactgt	aatgatctaa	aacacacgtc	tgctcatgag	180
acccgcaccc	cacccccgat	gctggggccg	ctcttggtat	ttcatgcctg	ctgccagcac	240
ccagggggag	ctccggaaat	gtctgctggg	ggctcggaat	acccaccttt	ctggtaaatgc	300
agcccagcgg	gtcccagcct	cgttttccag	ccctcactca	aaatggagtc	gctctggttc	360
gaacgcctct	gacaagtgtg	tacctacgtg	tcag			394

<210> 237

<211> 428

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(428)

<223> n = A,T,C or G

<400> 237

ttcggcagca	nnnaagaaga	ggccctcaga	gatctgacag	cctatgagtg	cgtggacacc	60
acctcagccc	actgagcagg	agtcacagca	cgaagaccaa	gcgcaaagcg	acccctgccc	120
tccatcctga	ctgctcctcc	taagagagat	ggcacccggc	agagcaggat	tctgccccct	180
tctgctgctt	ctgctgctgg	ggctgtgggt	ggcagagatc	ccagtcagtg	ccaagcccaa	240
gggcatgacc	tcatcacagt	ggtttaaaat	tcagcacatg	cagcccagcc	ctcaagcatg	300
caactcagcc	atgaaaaaca	ttaacaagca	cacaaaacgg	tgcaaagacc	tcaacacctt	360
cctgcacgag	ccttttctca	gtgtggccgc	cacctgccag	acccccaaaa	tagcctgcaa	420
gaatggcc						428

<210> 238

<211> 432

<212> DNA

<213> Homo sapiens

<400> 238

tctcatggag	gaacccatcc	attcgaattc	ggcacgagga	tcaactggct	atcatatctg	60
tttaatacat	ttactggagc	cagaaaccta	ggccatcatc	gaacgccagc	ccttgggtctg	120
agcctgcggc	tgtagatgtg	gaactcacag	catatgcatt	gttggcccag	cttaccaagc	180
ccagcctgac	tcacaaggag	atagcgaagg	ccactagcat	ataggcttgg	ttggccaagc	240
aacgcaatgc	atatgggggc	ttctcttcta	ctcacgatac	tgtagtgtgt	gtacaagctc	300
ttgccaaata	tgccactacc	gcctacgtgc	catctgagga	gatcaacctg	gttgtaaaat	360
ccactgagaa	tttccagcgc	acattcaaca	tacagccagc	taacagattg	gtatttcagc	420
aggataccct	gc					432

<210> 239

<211> 373

<212> DNA

<213> Homo sapiens

<400> 239

ggcacgaggc	aggacctcct	ctcccagatc	gccagctgc	aggaggagaa	caagcagctc	60
atgaccaacc	tctcccacaa	ggatgtcaac	ttctcagagg	aggagttcca	gaagcatgaa	120
ggcatgtcag	agcgggagcg	acagggtgatg	aacaagctga	aggaggtggg	ggacaaacaa	180
cgcgacgaga	tccgcgccaa	ggacagggag	ctgggcctga	aaaatgagga	cgttgaggct	240
ttacagcagc	agcagacacg	gctgatgaag	atcaaccatg	accttcggca	ccgggtcacg	300
gtggtggagg	cccaggggaa	agccctgatc	gaacagaagg	tggagctgga	ggcagacctg	360
cagaccaagg	agc					373

<210> 240

<211> 392

<212> DNA

<213> Homo sapiens

<400> 240

ggcacgagag	ctgaccgaga	tggacgtttt	ctacatcgcg	tcgcttgtgg	gccacgagtt	60
cgagcgggtc	attgaccagc	acgggtgtta	ggccatcgcg	cgcctcatgc	ccaaggctgt	120
gcgcgtgctg	gagatcttgg	aggtgctggg	cagtcgcctc	cacgtcgcgc	ccgagctgga	180
cgatctgcgc	ctggagcagg	acctcctctc	ccagatcgcc	cagctgctgg	aggagaacaa	240
gcagctcatg	accaacctct	cccacaagga	tgtcaacttc	tcagaggagg	agttccagaa	300
gcatgaaggc	atgtcagagc	gggagcgaca	ggtgatgaag	aagctgaagg	aggtggtgga	360
caaacaacgc	gacgagatcc	gcgccaagga	cg			392

<210> 241

<211> 434

<212> DNA

<213> Homo sapiens

<400> 241

gatcccatcc	attcgaattc	ggcacgagga	ttgattcacc	ttcacctgtg	ctgcactcca	60
gctgacccaa	gtaggaagcc	ggacgagctg	taaaacatga	acggaagagt	ggattatttg	120
gtcactgagg	aagagatcaa	tcttaccaga	gggccctcag	ggctgggctt	caacatcgtc	180
ggtgggacag	atcagcagta	tgtctccaac	gacagtggca	tctacgtcag	ccgcatcaaa	240
gaaaatgggg	ctgcggccct	ggatgggcgg	ctccaggagg	gtgataagat	cctttcggta	300
aatggccaag	acctaaagaa	cctgctgcac	caggatgctg	tagacctctt	tcgtaatgca	360
ggctatgctg	tgtctctgag	agtgcagcac	aggttacagg	tgcagaatgg	acctatagga	420
catcgagggtg	aagg					434

<210> 242

<211> 385

<212> DNA

<213> Homo sapiens

<400> 242

ggcacgagga	gagcgcggac	acctcctcaa	cccactgaac	aggagtcaca	gcacgatgac	60
cattcgcaaa	gcgacccctg	ccttccatcc	tgactgctcc	tcctaagaga	gatggcaccg	120

gccaaaacag	gattatgccc	ccttctgctg	cttctgctgc	tgccgctgag	tgtggcagag	180
atccccactca	gtgccaaacc	caagggcatg	acctcatcac	agtgggtttag	aattcagcac	240
atgcagccca	gccctcaagc	atgcaactca	gccatgaaaa	acattaacaa	gcacacaaaa	300
cgggtgcaaag	acctcaacac	cttcctgcac	gagcctttct	ccagtgtggc	cgccacctgc	360
cagacccccca	aaatagcctg	caaga				385

<210> 243

<211> 388

<212> DNA

<213> Homo sapiens

<400> 243

ggcacgagag	aaggcctgcg	gcaaagagat	gagcttattg	acaaacatgg	cttagttata	60
atccccgatg	gcactcccaa	tggtgatgtc	agtcatgaac	cagtggctgg	agccatcact	120
ggtgcgtctc	aggaagctgc	tcaggtcttg	gagtcaccag	gagaagggcc	attacatgtt	180
tggctacgaa	aacttgctgg	agagaaggaa	gaactactgt	cacagattac	aaaactgaag	240
cttcagttag	aggaggaacg	acagaaatgc	tccatgactg	atggcacagt	gggtgacctg	300
gcaggactgc	agaatggctc	agacttgcat	gtcatcgaaa	tgacagagaga	tgccaataga	360
caaattagcg	aatacaaatt	taagcttg				388

<210> 244

<211> 388

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(388)

<223> n = A,T,C or G

<400> 244

ggcacgaggt	cactgttgaa	gagttcaatc	ttaccagagg	gccctcaggg	ctgggcttca	60
acatcgccgg	tgggacagat	taccagtatg	tctccaacga	cagtggcatc	tacgtcagcc	120
gcatcaaaga	aaatggggct	gcggccctgg	atgggagggt	ccaggagggg	gataagatcc	180
tttcggtaaa	tggccaagac	ctaaagaacc	tgctgcacca	ggatgctgaa	cacctctttc	240
gtaatgcagg	ctatgctgtg	tctctgagag	tgacagacag	gttacaggcg	cagaatgtac	300
ctataggaca	tgcaggtgaa	ggggacccaa	gcggattccc	atattttattg	tgctgggtgcc	360
cngngctggc	ctctccctgg	tattcgcg				388

<210> 245

<211> 390

<212> DNA

<213> Homo sapiens

<400> 245

ggcacgaggc	tgtgtgtctc	ttttctcacc	ccagggcctg	gccatgtccc	ctttgggaag	60
cctgttccct	tacccttaca	cgtacatggc	cgcagcggcg	gccgcctcct	ctgcggcagc	120
ctccagctcg	gtgcaccgcc	accccttcct	caatctgaac	accatgcgcc	cgcggctgcg	180
ctacagcccc	tactccatcc	cgggtgccgg	cccggacggc	agcagtctgc	tcaccaccgc	240
cctgcctccc	atggcgggcg	ccgcggggcc	cctggacggc	aaagtgcgcc	ccctggccgc	300
cagcccggcc	tccgtggcag	aggactcggg	ctctgaactc	aacagacgct	cctccacgct	360
ctcctccagc	tccatgtcct	tgtcgcccag				390

<210> 246

<211> 397

<212> DNA

<213> Homo sapiens

<400> 246

ggcacgagac	cactgggacc	tcctgtctct	cgccatcctc	aacacagggc	tgtctctgtt	60
tgggtgcctt	tggatccatg	ccgcctaccc	ccactccccg	ctgcacgtgc	gagccctggc	120
cttagtgagg	gagcgtgtgg	agaacggaca	catctatgac	acgattgtga	acgtgaagga	180
gacgcggctg	acctcgctgg	gcgccagcgt	cctggtgggc	ctgtccctgt	tgctgctgcc	240

ggtcccgctt	cagtggatcc	ccaagcccgt	gctctatggc	ctcttcctct	acatcgcgct	300
cacctccctc	gatggcaacc	agctcgtcca	gcgcgtggcc	ctggtgggtc	aggaacacaa	360
ctggggaacc	ccccgacaca	ctacatcccc	gagggggg			397

<210> 247
 <211> 471
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(471)
 <223> n = A,T,C or G

<400> 247						
ttacggcgcg	tttggttagg	gacccaccg	attcgaattc	ggcacgagct	ctttttatatt	60
tcgctgatat	ctttctttta	ctaaatgcca	ccatccttac	ctgttcgggt	gtctgcgtgc	120
ctaatttttc	ctggctgtta	cacaagaacc	cggatttttag	ttgaactctg	gagcaaaaat	180
cctgcatcat	ttgtagggtg	gtgtcattgt	gactggctgc	tacctcccca	tgagtcttct	240
aaaataaaaac	ctgcaaattc	acatcttccc	catgcttcca	gagaatgcat	attcttcctt	300
tgaaaaaaga	aaacnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnggat	g	471

<210> 248
 <211> 403
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(403)
 <223> n = A,T,C or G

<400> 248						
ggcacgaggt	acagacatct	agttggcagg	agccaaagat	gttgccaaac	atgtagtann	60
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	120
nnnnnnnnnga	gtagaggatg	cctggtatga	ggcaatattt	gggataggga	aggggaagctt	180
gggatttttag	ctacgtagag	acacttgaaa	attggaggga	ggaaaggagt	gggtggcttt	240
ggagatgttc	tggaatatgt	gaatgagggg	agtggagggg	ncctgnnngc	tctgnggaag	300
gccangcccg	gtttcctgtc	tttcancctc	ttccaggaaa	attacgggca	gaaagaggct	360
gagaaagtgg	tcccggggaa	ggcgctttat	gaagagcttg	gtg		403

<210> 249
 <211> 316
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(316)
 <223> n = A,T,C or G

<400> 249						
ccgcttaaag	gcgccttctt	ttaatgcaat	cattttgaac	atgtgcgaca	gtcgagaata	60
ctaattggat	caatcttgat	atactctacc	taaagacagt	ctagaaacct	gggggagaaa	120
gaactcacgg	cacaaaacat	tgggccgaga	acggaattct	ctgtaagcct	agttgctgaa	180
acttctgtct	gtaaccagaa	gccagtttta	tctatcggtc	actgaaacac	ccactgtgtg	240
ttgctcactc	cctcactcac	cgaacanaac	ctgctacctc	cgcatgaatc	tactagtgcc	300
gataaactat	atcaga					316

<210> 250
 <211> 419

<212> DNA

<213> Homo sapiens

<400> 250

ggcagcgagat	atcagtcgaag	ggctcttcaa	gacacagcag	aaacctcacc	gggcctcggg	60
ctgcctccca	ctgggtccca	tggccaccac	cttgaccttg	gaaagctctg	ttatatggaa	120
ggtagggagg	acactatttc	cctcaactac	ttctagtaaa	aagctcagtt	ctctcccag	180
cagcaagagg	gcacctgtga	acacctgagt	cacagcgcat	tcctcctctg	cttagaacat	240
tcgatggctc	ccaccttact	tgcagtaa	gctgaggtcc	ttcctgtggc	ccccggggcc	300
ctgcatgata	tgatccatcc	cttacctacc	ctcatctctc	cactggcctc	cccacacttg	360
ctccccctcg	gacactctgg	actacttgct	gctatctgaa	cataccaggc	ccctgcccc	419

<210> 251

<211> 434

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(434)

<223> n = A,T,C or G

<400> 251

ggcagcaggg	ggcctccacc	ggtgactcgg	gcctggattc	cacggccatg	gcctctgccg	60
ctgcggcgca	gggactgtcc	ggggcgctcc	cggacaccct	gcccttccac	ctccagcagc	120
acgtcctggc	ctctcagggc	ctggccatgt	cccctttcgg	aagcctgttc	ccttaacct	180
acacgtacat	ggacgcagcg	gcggccgcct	cctctgcggc	agcctccagc	tcgggtgcacc	240
gccacccctt	cctcaatctg	aacaccatgc	gcccgcggct	gcgctacagc	ccctactcca	300
tcccgggtgcc	ggtcccggac	ggcagcagac	tgctcaccac	cgccctgccc	tccatggcgg	360
cggccgcggg	gcccctggac	ggcaaagacg	cgcctctggc	cgccagcccc	gcctcggagg	420
cagtggactc	ggcg					434

<210> 252

<211> 425

<212> DNA

<213> Homo sapiens

<400> 252

ggcagcagaa	agcactcagc	ctggggaatg	aactctgcc	caatgatgat	ggctgtgacc	60
actccccgca	gagagttctt	gaagaggagc	tcggcaggga	ctggcaggcc	aaggtggcct	120
ccttgaggga	ggtgcccttt	gccgctgcct	caattgggca	ggtgcaccag	ggcctgctga	180
gggacgggac	ggaggtggcc	gtgaagatcc	aggtgagagg	ggaggctggg	cagggtaggg	240
gcgggcaccc	tgctagccca	gagaagtgc	tcccaccttc	tctccctccc	ttctcccttt	300
acagtacccc	ggcatagccc	agagcattca	gagcgatgtc	cagaacctgc	tggcgggtact	360
caagatgagc	gcggccctgc	ccgcgggcct	gtttgccgag	cagagcctgc	aggccttgca	420
gcagg						425

<210> 253

<211> 395

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(395)

<223> n = A,T,C or G

<400> 253

ggcagcagca	gacatctagt	tggcaggagc	caaagatgtt	gccaaacatg	tagtannnnn	60
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	120
nnnnngagta	gaggatgcct	ggtatgaggc	aatatttggg	atagggaagg	gaagcttggg	180
attttagcta	cgtagagaca	cttgaaaatt	ggagggagga	aaggagtggg	tggctttgga	240
gatgtttctg	aatatgtgaa	tgaggggagt	ggaggggtcc	tggaggtctt	ggggaaggcc	300

aagcccgttt tcctgtcttt caacctcttc caggaaaatt acgggcagaa ggaggctgag 360
 aaagtggccc ggggtgaatgc gctatatgac gagct 395

<210> 254
 <211> 307
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(307)
 <223> n = A,T,C or G

<400> 254
 agtcgtcttc ttttaaatgta atcattttga acatgtgtga aagttgatca tacgaattgg 60
 atcaatcttg aaataactcaa ccaaaagaca gtcgagaagc caggggggaga aagaactcag 120
 ggcacaaaat attgggtctga gaatggaatt ctctgtaagc ctagttgctg aaatttcctg 180
 ctgtaaccag aagccagttt tatctaacgg ctactgaaac acccactgtg ttttgctcac 240
 tccctcactc accgatcaaa acctgctacc tccccaagac tttactagtg ccgataaact 300
 ttctcan 307

<210> 255
 <211> 312
 <212> DNA
 <213> Homo sapiens

<400> 255
 agtcgtcttc ttttaaatgta atcattttga acatgtgtga aagttgatca tacgaattgg 60
 atcaatcttg aaataactcaa ccaaaagaca gtcgagaagc caggggggaga aagaactcag 120
 ggcacaaaat attgggtctga gaatggaatt ctctgtaagc ctagttgctg aaatttcctg 180
 ctgtaaccag aagccagttt tatctaacgg ctactgaaac acccactgtg ttttgctcac 240
 tccctcactc accgatcaaa acctgctacc tccccaagac tttactagtg ccgataaact 300
 ttctcaaaga gc 312

<210> 256
 <211> 415
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(415)
 <223> n = A,T,C or G

<400> 256
 ggcacgagca ggagcagctg gcaaggaggaga aggacacggt gaagatgctg caggaacagc 60
 tggaaaaggc agcgcgtgcc tggcgccaaa gcagggcggg aggagtcgag ctgccgggag 120
 ccccggggag gcaggaccgg gagaggcaga gctgggcgga gtcgtcaagc tgctgggagc 180
 gctgggctgg gagccccagg ggaggcagag ctgggcggag gtagtgggga cagagacttc 240
 ctaacgaggg cttcagccca cccggccac caccaccct tctgggggtc ccttgctggg 300
 aagcgagtgt ctgatcccc tgctggccca ggtcctcact ttgcacctgt gtgggcccct 360
 tagccagtgc tccagcccct gccctgcagg atgatggttt cccctcagct cccan 415

<210> 257
 <211> 396
 <212> DNA
 <213> Homo sapiens

<400> 257
 agaaagggtg agtgaggtgc tgtcctgggg ttctccaagt ttgagagcat ggatgcatgt 60
 ggtttgaagc tgaagtgggc ctgggggaat gggttgaagg cagaagcaac cagtttggag 120
 ggaaggcatt tggatatcca gccctttctc tgtggccttg gccctgggtc tgtcctgtta 180
 cccccacca tacctgtctg ctgcgcactc tgtgcttctg tagcattctc gcttctggcc 240

tttaaagtgtg	gcaaggggag	gttaataagc	acctaggtgg	ctgagtgtct	ctgtcttctg	300
gcttggttcac	aggacttcga	gtaagaaggt	gatttacagc	cagcctagtg	cccgaagtga	360
aggagaattc	aaacagacct	cgtcattcct	ggtgtg			396

<210> 258
 <211> 431
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(431)
 <223> n = A,T,C or G

<400> 258						
gnnggagggc	ctgcggcaaa	gagatgagct	tattgagaaa	catggcttag	ttataatccc	60
cgatggcact	cccaatggtg	atgtcagtc	tgaaccagtg	gctggagcca	tactgttgt	120
gtctcaggaa	gctgctcagg	tcttgagtc	agcaggagaa	gggccattag	atgtaaggct	180
acgaaaactt	gctggagaga	aggaagaact	actgtcacag	attagaaaac	tgaagcttca	240
gttagaggag	gaacgcacaga	aatgctccag	gaatgatggc	acagtgggtg	acctggcagg	300
actgcagaat	ggctcagact	tgcagttcat	cgaaatgcag	agagatgcca	atagacaaat	360
tagcgaatac	aaattttaagc	tttcaaaagc	agaacaggat	ataactacct	tggagcaaag	420
tattagccgg	c					431

<210> 259
 <211> 404
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(404)
 <223> n = A,T,C or G

<400> 259						
ggcacgagca	ggagcagctg	gcaagggaga	aggacacggt	gaagaagctg	caggaacagc	60
tggaaaaggc	agcgcggtgcc	tggcgccaaa	gcagggcggg	aggagtcgag	ctgccgggag	120
ccccggggag	gcaggaccgg	gagaggcaga	gctgggcgga	gtcgtaagc	tgctgggagc	180
gctgggctgg	gagccccagg	ggaggcagag	ctgggcggag	gtagtgggga	cagagacttc	240
ctaacgaggg	cttcagccca	cccggccccc	caccacacct	tctgggggttc	ccttgctggg	300
aagcgagtgt	ctgatcccc	tgctggccca	ggctcctcact	ttgcacctgt	gtggggccct	360
tagccagtgc	tccagcccct	gccctgcagg	atgatggttt	cccn		404

<210> 260
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 260						
ggcacgagat	ctccctgcct	tgtgagcagc	tggccggcgg	ctctgggaca	ggcggggatg	60
ggagggagtc	taccgggcca	ctgtagagct	ggtagctggg	agctggagct	gtagagttcc	120
aggctgggag	ctggagagcc	ctgggtgaga	gggaggccta	taggggcccc	gggggacaca	180
ccaggcttga	gggtagtagg	tgctggaggc	agagcctggc	ctgtccaggg	tgggacctca	240
cgacccaccc	tgtccggccc	ccagctcgga	ggagcttcta	cgtgtatgcg	ggcatcctgg	300
cactgctcaa	cctactgcag	gggctgggga	gtgagctgct	gtgcttcgac	atcatcgagg	360
ggctctgggtg	cgtggggggc	gcagggagtc	tgectcgtgg	gg		402

<210> 261
 <211> 402
 <212> DNA
 <213> Homo sapiens

<400> 261

```

ggcacgagat ctccctgcct tgtgagcagc tggccggcgg ctctgggaca ggcggggatg      60
ggagggagtc taccgggcca ctgtagagct ggtagctggg agctggagct gtagagttcc      120
aggctgggag ctggagagcc ctgggtgaga gggaggccta gaggggcccc gggggacaca      180
ccaggcttga gggtagtagg tgctggaggc agagcctggc ctgtccaggg tgggacctca      240
cgaccacccc tgtccggccc ccagctcggg ggagcttcta cgtgtatgcg ggcacccctgg      300
cactgctcaa cctactgcag gggctgggga gtgtgctgct gtgcttcgac atcatcgagg      360
ggctctgggtg cgtggggggcc gcagggtgtc tgccctcgtgg gg                      402

```

<210> 262

<211> 151

<212> DNA

<213> Homo sapiens

<400> 262

```

gccgaatatg aagctacgtc cgggtatccg ggttcctctgt aattgctttc tgatccctgg      60
tacttagatt tgattaccta tggaccacat tggtagaact actatatggg ggaacctcct      120
gattttggggc ggtctcaaaa acaaaaaaaaaa c                      151

```

<210> 263

<211> 404

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(404)

<223> n = A,T,C or G

<400> 263

```

ggcacgaggg aacgtggaag gactagactg cctgagtctt ctgannnnnn nnnnnnnnnn      60
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnntt      120
ggacctaccc cccgagtggg ttgccagggg ctctcaggcc ttcggctaca gactgagggc      180
tgcattatca gctttcctac ttttgagggt ttgggacttt actggctttc ttgctcctca      240
acttgcagat ggcctgttgt gggacctcac ctgtgtatca tgtacatgag ggaaatacac      300
accctccca gggatgatgg aagggttaagg tcctaacacc tcctgcacat ctgagcagct      360
gcacattgaa ccagatagtc ctggaatgtg ggaaaacaga ggcn                      404

```

<210> 264

<211> 380

<212> DNA

<213> Homo sapiens

<400> 264

```

ggcacgaggg gaacgggaag ccgggaccca gaactcttgt ctttcaggat aaagtggcca      60
gggtgtacga agccccgggc tttttcctgg acctggagcc catcccggga gccttggaag      120
ctgtgcggga gatgaacgac ctaccggaca cgcaggctct catctgcacc agccccctgc      180
tgaagtacca ccactgtgtg ggtgagaagt accgctgggt ggagcagcac ctggggcccc      240
agttcgtaga acgaattatc ctgacaaggg acaagacggt ggtcttgggg gacctgctca      300
ttgatgacaa ggacacagct cgaggccagg aggagacccc aagctgggag cacatcttgt      360
tcacctgctg ccacaatcgg                                     380

```

<210> 265

<211> 440

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(440)

<223> n = A,T,C or G

<400> 265

```

ggcagaggcg tggacaccac ctacgcccac tgagcaggag tcacagcacg aagaccaagc      60

```

gcaaagcgac	ccctgccctc	catcctgact	gctcctccta	agagagatgg	caccggccag	120
agcaggattc	tgcccccttc	tgctgcttct	gctgctgggg	ctgtgggtgg	cagagatccc	180
agtcagtgcc	aagcccaagg	gcatgacctc	atcacagtgg	tttaaaattc	agcacatgca	240
gcccgacct	caagcatgca	actcagccat	gaaaaacatt	aacaagcaca	caaaacgggtg	300
caaagacctc	aacaccttcc	tgcacgagcc	tttctccagt	gtggccgcca	cctgccagac	360
ccccaaaata	gcctgcaaga	atggcgataa	aaactgccac	caaagccacg	ggcccgtggt	420
cctgaccatg	tgaagctccn					440

<210> 266

<211> 396

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(396)

<223> n = A,T,C or G

<400> 266

gcacgaggag	gaacgtggaa	ggactagact	gcctgagtct	tctgannnnn	nnnnnnnnnn	60
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnt	120
tggacctacc	ccccgagtgg	tttgccaggg	gctctcaggc	cttcgggtac	agactgaggg	180
ctgcattatc	agctttccta	cttttgaggt	tttgggactt	tactggcttt	cttgctcctc	240
aacttgcaga	tggcctgttg	tgggacctca	ccttgtgatc	atgtacatga	gggaaataca	300
cacccctccc	agggatgatg	gaagggttaag	gtcctaacac	ctcctgcaca	tctgagcagc	360
tgcacattga	accagatagt	cctggaatgt	gggaac			396

<210> 267

<211> 429

<212> DNA

<213> Homo sapiens

<400> 267

ggcacgagga	tctgacagcc	taggagtgcg	tggacaccac	ctcagcccac	tgagcaggag	60
tcacagcacg	aagaccaagc	gcaaagcgac	ccctgccctc	catcctgact	gctcctccta	120
agagagatgg	caccggccag	agcaggattc	tgcccccttc	tgctgcttct	gctgctgggg	180
ctgtgggtgg	cagagatccc	agtcagtgcc	aagcccaagg	gcatgacctc	atcacagtgg	240
tttaaaattc	agcacatgca	gcccgacct	caagcatgca	actcagccat	gaaaaacatt	300
aacaagcaca	caaaacgggtg	caaagacctc	aacaccttcc	tgcacgagcc	tttctccagt	360
gtggccgcca	cctgccagac	ccccaaaata	gcctgcaaga	atggcgataa	aaactgccac	420
cagagccac						429

<210> 268

<211> 405

<212> DNA

<213> Homo sapiens

<400> 268

ggcacgaggc	ggcttccttg	cccgcgagca	gtaccgcgcc	ctgcggcccg	acctggcgga	60
taaagtggcc	agtgtgtacg	aagccccggg	ctttttcctg	gacctggagc	ccatcccggg	120
agccttggac	gctgtgcggg	agatgaacga	cctaccggac	acgcaggtct	tcactctgcac	180
cagccccctg	ctgaagtacc	accactgtgt	gggtgagaag	taccgctggg	tggagcagca	240
cctggggccc	cagttcgtag	aacgaattat	cctgacaagg	gacaagacgg	tgggtcttggg	300
ggacctgctc	attgatgaca	aggacacagt	tcgaggccag	gaggagaccc	caagctggga	360
gcacatcttg	ttcacctgct	gccacaatcg	gcacctggcc	tgccc		405

<210> 269

<211> 372

<212> DNA

<213> Homo sapiens

<400> 269

ggcacgagaa	ccctgaggcc	tggctatggt	accaccgggt	ggtaggtgcc	cagcgctgcc	60
------------	------------	------------	------------	------------	------------	----

ccatcgtgga	caccttcttg	caaacagaga	caggtggcca	catgttgact	ccccttcctg	120
gtgccacacc	catgaaaccc	ggttctgcta	ctttcccatt	ctttggtgta	gctcctgcaa	180
tcctgaatga	gtccggggaa	gagttggaag	gcgaagctga	aggttatctg	gctgccagcg	240
ggaccaggat	ggctattact	ggatcactgg	caggattgat	gacatgctca	atgtatctgg	300
acacctgctg	agtacagcag	aggtggagtc	agcacttggt	gaacatgagg	ctgttgcata	360
ggcacctgtg	gg					372

<210> 270

<211> 411

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(411)

<223> n = A,T,C or G

<400> 270

ggcacgagag	ctctcggcgc	acggcccagc	ttccttcaaa	atgtctactg	ttcacgaaat	60
cctgtgcaag	ctcagcttgg	agggtgattg	tccaggaagt	tattccagat	gaagacttat	120
acnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	180
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	411

<210> 271

<211> 302

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(302)

<223> n = A,T,C or G

<400> 271

ctgagtgtga	cactcagaga	gtgtgttata	tagacagaga	gagagcgcgc	gcctctgtcc	60
cccccttgtg	gtgtgccccca	ctccagtgcg	cccagatccg	tgcccccccc	cggagcgcgc	120
tgctccctnn	nnnnnnnnnag	tgtgcacacc	cccctcccc	tctcatgagt	gccacatat	180
atattcctgt	gtgacccctc	ccccccctg	ccagtcagt	tccccgcgcg	agcgcgagtc	240
actgttttat	tttttctcgc	ccccaagaag	ggatagcgat	gtgtctctcc	cctcctccca	300
ca						302

<210> 272

<211> 429

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(429)

<223> n = A,T,C or G

<400> 272

ggcacgagat	gtggtacaga	catctagttg	gcaggagcca	aagatgttgc	caaacatgta	60
gtannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	120
nnnnnnnnnn	nnngagtaga	ggatgcctgg	tatgaggcaa	tatttgggat	agggaaggga	180
agcttgggat	tttagctacg	tagagacact	tgaaaattgg	agggaggaaa	ggagtgggtg	240
gctttggaga	tgttctggaa	tatgtgaatg	aggggaagt	gaggggcctg	gaggctctgg	300
ggaaggccaa	gcccgttttc	ctgtctttca	acctcttcca	ggaaaattac	gggcagaagg	360
aggctgagaa	agtggcccgg	gtgaaggcgc	tatatgagga	gctggaactg	tcaacagtgg	420
tcttgcaaaa						429

<210> 273
 <211> 471
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(471)
 <223> n = A,T,C or G

```
<400> 273
tgtttgtgcat ttgggcatcc caccgattcg aattcggcac gagaaagcat tgaagagacc      60
tcaaggctttt aagaaatgag taggccaaaa tctaagtcaa aggagaatct gtactggggc      120
cccccgctgcc ctgagggtcat tggccaagcc aagccgaacc tgagctttga tcctgatggt      180
ttgggggagtg aggaagacag aagtggaagc ccagttctca cccaagagg ggacacaaat      240
ggatgaccct cccatgatgc tgagaccca aaaggctaca cactcaagct aaaagccaga      300
ggaaatccca tcctgccacc cacaagactt caaggaaagt tgttttggtg ctgagcagag      360
caggggaaga aggaaaacag cccttaagga gctccagcca ctggccagcc ttcatgtgac      420
tctagcccaa attcattccc atcacctggg gtggaagggc cagaaatctc n      471
```

<210> 274
 <211> 391
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(391)
 <223> n = A,T,C or G

```
<400> 274
ggcagcaggt aaactctcta taagtgttca gtgttgacat agcctttgtg catagnnnnn      60
nnnnnnnnnn nnnnnnnnnn nnnnnnnnttt tttgccccac ctggaaaaaa ggggatgcnn      120
nnnnnnnnngg ggggaaaaaac aattcttaag ggcccttttg ccataaactt ttttccgggc      180
cacctttgtt acttttggtc ctggaagggg tttttttggg gggcccacgg ggagggggccc      240
cataggtaaa ctcggaaaac tttttctaac ccgggttagt gtttttaatt aaaacaaaaa      300
annnnnnnnn nnttggaatc cttttcttta aaaaaattaa tctctcaaag gaaaacaaag      360
nnnnnnnnnn ngggggggccc ctttcgttta g      391
```

<210> 275
 <211> 339
 <212> DNA
 <213> Homo sapiens

```
<400> 275
cactccgggg gctctatttg tgtgctctgc acccagtttt ttatacactc cacgcttttg      60
atataacatc tagcgccacg gtgcctatgt gtacacaccc tctctctata tatagatacc      120
tctgtgcgca catatagagg ggaaaagaga gatatatcta ttatatatac atttctacac      180
aactgtctct gggggggtcag agaacgcgcg caccctctc ttttgagaga aggagactct      240
gtccccctc tctggggcgc agggaggccc catggcatga agaaaaatac tcacttatat      300
ctctctctct cactctctgt ttgcgaaaaa acacacagg      339
```

<210> 276
 <211> 434
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(434)
 <223> n = A,T,C or G

<400> 276

ccctagctac	ttgctctttg	tgcaggatgc	catagattcg	tgggctcctg	ccttttctca	60
accccagagt	gcctgaccag	ttctaccgcc	tgtggctatc	cctcttcctg	cacgccggga	120
tcttgcactg	cctgggtgtcc	atctgcttcc	agatgactgt	cctgcgggac	ctggagaagc	180
tggcaggctg	gcaccgcata	gccatcatct	acctgctgag	tgggtgtcac	ggcaacctgg	240
ccagtgccat	cttcctgcca	taccgagcag	aggtgggtcc	tgctggctcc	cagttcggca	300
tcttgacctg	cctcttcgtg	gagctcttcc	agagctggca	gatcctggcg	cggcctggc	360
gtgccttctt	caagctgctg	gctggggagg	cttttctctt	cacctttggg	ctgctgccgt	420
ggattgacaa	cttn					434

<210> 277

<211> 378

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(378)

<223> n = A,T,C or G

<400> 277

ggcacgagaa	aaagtaccgc	tccagagcag	gagcctaggg	agccgagagg	gtgcccgaac	60
ctgagtctga	gttgccggcca	cttcaggagc	tgagaggagc	aggatggaac	tgaggatcc	120
aaagatgaat	ggagccctcc	cttcggatgc	tgtgggctac	aggcaagaac	gtgagggctt	180
cctgcccagt	cgtggctcctg	ctcctgggag	caagccggtc	cagttcatgg	atttcgaggg	240
gaagacatcg	tttggaatgt	cagtgttcaa	cctcagcaac	gccatcatgg	gcagcggcat	300
cctggggctg	gcctatgcca	tgggccacac	gggggtcatt	ttctttctgg	gcctgctgct	360
gngccatgcg	cttctgcc					378

<210> 278

<211> 302

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(302)

<223> n = A,T,C or G

<400> 278

ccccnctct	cgcnnnnnn	nnnnncgttt	tcactcccgg	gagtcccctt	gtttttggcc	60
cggatccggg	ttctttcttt	cccgtggtgc	cgcggtttgg	agtgttttat	cttttcttca	120
catggggggc	tggggagttc	cccagaacct	ccagggggaa	acccccctcc	tatgaaaatg	180
acacatgagc	ccctccttcc	ggtggcgggg	acctgtctct	ctaagacctt	tttctgggaa	240
aggggtcttt	gtttgtatga	ccccaccgac	gcggggggct	ttctatgggc	cgcccccccc	300
cg						302

<210> 279

<211> 405

<212> DNA

<213> Homo sapiens

<400> 279

ggcacgaggc	ctcattggag	acattgacaa	tgccatgagg	accttcctca	actactacac	60
tgtatggaag	cagtttgggg	ggctcccggg	attctacaac	attcctcagg	gatacacagc	120
ggagaagcga	gagggctacc	cacttcggcc	agaacttatt	gaaagcgcaa	tgtacctcta	180
ccgtgccacg	ggggatccca	ccctcctaga	actcggaaga	gatgctgtgg	aatccattga	240
aaaaatcacc	aaggtggagt	gcggatttgc	aacaatcaaa	gatctgcgag	accacacgct	300
ggacaaccgc	atggagtcgt	tcttcctggc	cgagactgtg	aaataacctt	acctcctggt	360
tgaccaaac	aacttcatcc	acaacaatgg	gtgcaccttc	gacgc		405

<210> 280

<211> 415

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(415)

<223> n = A,T,C or G

<400> 280

ggcagcaggg	tcacctgtgc	tgcccctcct	taatctcgta	tgatgggtcac	agtccgggtgg	60
ccgtgggggt	gctctgcctt	ccctgggtccc	cactgcccac	atctgtggac	tgccccttcc	120
aaagaccctt	gggggggggt	ggananattc	aatcttacca	aactcaacga	tccatccatt	180
tcattgttact	gatattacat	gcggacaccc	ctggatcata	ttattcaa	ccagtcatt	240
attctgcatt	catgaccttt	tgataactcc	atcatgacct	acttgacggt	cactgacct	300
gcttactgga	ttccgccttg	taacaataaa	atctatttaa	actnnnnnnn	nnnnnnnnnn	360
nnnaccagcc	cacataaaat	atgattgaat	caatttctta	taccttcact	agaat	415

<210> 281

<211> 389

<212> DNA

<213> Homo sapiens

<400> 281

ggcagcaggt	agactggggg	ctcactgatt	gcattgacac	ttttcatcat	gggtccccgg	60
gggctcacgt	ggagtctgac	acatgaatac	atggctatca	tgtctgtcac	cttcaatggg	120
gaaaacaaac	tttgtaatgg	taggaaacac	aacagggtaca	ataatttaca	aaaatatgtt	180
tgccacattt	cagggcaagg	caaaatgcag	aggagacata	tgttaaaatc	ttatcattca	240
catttgttct	ttttatcttt	aagatgaagc	tcttacacca	agtgtcacga	gtctggagaa	300
cagatgggtt	gaagagctgt	tcttataaaa	taagatctgg	ggaacacaat	cctttatata	360
tcaacatcac	agtggatttt	tggattggg				389

<210> 282

<211> 371

<212> DNA

<213> Homo sapiens

<400> 282

ggcagcagat	agaatccgag	gcattgatat	cattaaatgg	atggagcgct	accttaggga	60
taagaccgtg	atgataatcg	tagcaatcag	ccccaaatac	aaacaggacg	tggagggcgc	120
tgagtgcgag	ctggacgagg	atgagcatgg	cttacatact	aagtacattc	atcgaatgat	180
gcagattgag	ttcataaaac	aaggaagcat	gaatttcaga	ttcatccctg	tgtcttccc	240
aaatgctaag	aaggagcatg	tgcccacctg	gcttcagaac	actcatgtct	acagctggcc	300
caagaataaa	aaaaacatcc	tgctgctggc	gctgagagag	gaagagtatg	tggctcctcc	360
acgggggcct	c					371

<210> 283

<211> 413

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(413)

<223> n = A,T,C or G

<400> 283

ggcagcaggt	gggagacacc	acttgtcttt	atgtgggtct	caaagatgat	gtagaatttc	60
ctttaatttc	tcgcagtctt	cctggaaaaat	attttccttt	gagcagcaaa	tcttgtaggg	120
atatcagtga	aggtctctcc	ctccctcctc	tcctgnnnnn	nnnnnnngga	aacaaagt	180
tgcttttgtt	ccccagcctg	aaggggaagg	gctcaatttt	gggttaaccaa	aaccttgcc	240
tccgggggta	aagcaattct	ccggcctaac	cctttggaga	acctggggtta	ataggcgcag	300
gccccagggc	cgggttaatt	ttgggtttta	agaaaaaaca	gggtttctca	atgtggggca	360
ggcgtggcca	aaacccccac	cctaagggga	tcggccctcc	ttggcctccc	aan	413

<210> 284
 <211> 409
 <212> DNA
 <213> Homo sapiens

```
<400> 284
ggcacgaggc ctgggggatgc tccctgctaa gtgggcctgc tcccaccctt gccataaagc      60
tctgaggcag cctgagcctg cctggggggc cccactgtga ccctgccgca gtcttcctgg      120
gtccctgcgt cctcttaagg ggcagtgaca cctgcctcgc tggccctgtg tgggtggcag      180
gccccactgt ttgggatatc acatggccag gcacgtggtg agcctgtcca gggcggcacgc      240
ctgcaggcgc gtgctcggtc acacactgcc ttgtgtggcc ctccctgtccg gtgcagcctg      300
gacctggacg cctggatcaa tgagccactc tcggacagcg agtcagagga cgagaggccc      360
agggccgtct tccacgagga ggagcagcgg cgtcccaagc accggccgt      409
```

<210> 285
 <211> 404
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(404)
 <223> n = A,T,C or G

```
<400> 285
ggcacgagcc acttcacccc cttgggggct gcttattcac tctggggatt cgccatggac      60
acgtctcaac tgcgcaagct gctgcccctg tttccctgcc cctccagatt gcctggagat      120
ctattttgtt tccttttgtg tttctttttc tgttttgagt gtctttcttt gcagggtttct      180
gtagccggaa gatctccgtt ccgctcccag cggctccagt gtaaattccc cttccccctg      240
gggaaatgca ctaccttgtt ttgggggggt taggggtgtt tttgtttttc agnngntttg      300
nttttttggg nnnnnnnnnn gntttgactt ttttnnctt tattttggag ggtaatggaa      360
agaataggaa aatcaggcag gggggagaat ggttgtttat tctt      404
```

<210> 286
 <211> 441
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(441)
 <223> n = A,T,C or G

```
<400> 286
ggcacgaggg aagcgggtgg tgtgtgtccc ctgtttactt ttagctgagc tgggggttggg      60
tgtacgggtt ctgttcctct gagccctgcg gccacactga tgtttacgtg tgtgtgtgag      120
ggggggcggc gctncncnnn cccccccan nggcctctat ccttgtgaag ctctcctcaa      180
tctaatactt attgcccctg actccaaatc ttccaccttt tgccctttat tatactctatg      240
ttcattacct taggtcagct gttctctatt atgacactga ttcatacttt tgttttttga      300
taagtactta tttcctctct cattgttgct aatatcctct tccttttttc ctttgtctac      360
tctcacttca tctataaaac tcttacatat ctctccacta atttctttga actaacaatt      420
tttatataga atttaagcct g      441
```

<210> 287
 <211> 387
 <212> DNA
 <213> Homo sapiens

```
<400> 287
ggcacgagca gccctggaat tccgcaagca cccggaggcc ggggggtctc cgcgggcgtc      60
ccatgcggag gacatggtgc gccgtgtact cttccccacg aacctaggga ccggtccccc      120
cgccggaaact gcttcctacc tggtcgggtc ccggcagctg aatctggcca gcccaacctc      180
ccggtcgcta tggcaccac aggcctaaca ttccgcgagtc caccttccgc cgtccgcgag      240
```

gaaaacctga	ttggcgccct	cttggcgatc	ttcggggcacc	tcgtgggtcag	cattgcactt	300
aacctccaga	agtactgcc	catccgcctg	gcaggctcca	aagatccccg	ggccttatatt	360
aaagacccaaa	actggtggct	tgggcct				387

<210> 288
 <211> 439
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(439)
 <223> n = A,T,C or G

<400> 288						
ggcacgagg	aggctggaag	cgggtggtgt	gtgtcccctg	tttactttta	gctgagctgg	60
ggttggtgt	acgggttctg	tccctctgag	ccctgcggcc	cacctgatgt	ttacgtgtgt	120
gtgtgaggg	gggcggcggn	nncannnnnn	nnnnnnngan	tctttttcca	ataacaatat	180
taattaatcc	aatctttttt	cttcttctct	tctttctact	cttttttctc	cttttttttt	240
atttactttt	actcatcctc	ctttcttcat	ttactctgtc	ttttgtatta	ctagcttctt	300
ctctttcgca	attttccttt	attgttgtca	ctcttttggg	aataacgtac	tcttatgaga	360
agttgtttcc	tctttattta	catttggttg	tcttctcctt	tcataattta	ttttacgtat	420
gtttgtggag	ttttttctt					439

<210> 289
 <211> 170
 <212> DNA
 <213> Homo sapiens

<400> 289						
atgagtggtc	ttttaattag	gaacaaatct	aatggaaagg	agagttgact	gaagttggcc	60
cacaggattg	tgagctgggc	agagccttca	tgaaggcttg	ccaccttggg	acgcccgaatt	120
taatgggggg	gcctgctgta	aggcaaaagg	cttttttgga	aattgctggg		170

<210> 290
 <211> 393
 <212> DNA
 <213> Homo sapiens

<400> 290						
ggcacgagg	agactggggg	ctcactgagt	gcagtgcac	ttttcatcat	gggtccccgg	60
gttctcacgt	ggagtctgac	acatgaatac	atggctatca	tgtctgtcac	cttcaatggg	120
gaaaacaaac	tttgtaatgg	taggaaacac	aacaggtaca	ataattttaca	aaaatatggt	180
tgccacattt	cagggcaagg	caaaatgcag	tgagacata	tgttaaattc	ttatcattca	240
catttgatct	ttttatcttt	aggatgaagc	tcttacacca	agtgtcacga	gtctggagaa	300
cagatggggg	gagtagttgt	tcttataaat	tagtatctgt	ggaacacaa	cctttatata	360
tcaacatcac	agtggatttc	tggcttggtg	cat			393

<210> 291
 <211> 430
 <212> DNA
 <213> Homo sapiens

<400> 291						
ggcacgagg	atagaatccg	aggcattgat	atcattaaat	ggatggagcg	ctaccttagg	60
gataagaccg	tgatgataat	cgtagcaatc	agccccaaat	acaaacagga	cgtggaaggc	120
gctgagtcgc	agctggacga	ggatgagcat	ggcttacata	ctaagtacat	tcatcgaatg	180
atgcagattg	agttcataaaa	acaagggaagc	atgaatttca	gattcatccc	tgtgctcttc	240
ccaaatgcta	agaaggagca	tgtgccacc	tggcttcaga	acactcatgt	ctacagctgg	300
cccaagaata	aaaaaaacat	cctgctgcgg	ctgctgagaa	aagaaaaaga	tgtggctcct	360
tcacgggggc	ctcttgccac	ccttcaagtg	ggtcccttgt	gacacccgtc	aatcccagat	420
cactgaggcc						430

<210> 292
 <211> 423
 <212> DNA
 <213> Homo sapiens

<400> 292
 atcccatcga ttccgaattcg gcacgagggga agcaaggggca cccgccttat ggatggaatt 60
 gaggggaagg caccgggggc tcctgcatcg agcttccctc ctatattcaa tgaggaaatg 120
 accctgcaga aggctggctg cagatgcccc tgcctcccg ctttgctgc ttggagtttg 180
 atggacacgt ggtcctgtca gggctacagc aggtctatgg tctttggtaa cggaaagcgc 240
 tgggtgaaaca gtgagctttc ccgtgggtgc ttttccctga cgccaacaac cagggcaagc 300
 tgcctgtcct gctgcttggc cgctcctcag agctgcggcc gggagagttc gtggtcgcca 360
 tcggaagccc gttttccctt caaaacacag tcaccaccgg gatcgtgagc accaccagc 420
 gag 423

<210> 293
 <211> 409
 <212> DNA
 <213> Homo sapiens

<400> 293
 ggcacgaggc taggagtact ggcctagatg gttatagaag tccatgccag gaggtcgtct 60
 gcagtcagag ggtggttctg ggctggactc cagccccttc ctgtcggagg ccaatgccga 120
 gcggattgtg cagaccttat gtacagtctg aggggccgcc ctcaaggttg gccagatgct 180
 cagcatccag gacaacagct tcatcagccc tcagctgcag cgcattcttg agcgggtccg 240
 ccagagcgcc gacttcatgc cccgctggca gatgctgaga gttcttgaag aggagctcgg 300
 cagggactgg caggccaagg tggcctcctt ggaggaggtg ccctttgccg ctgcctcaat 360
 tgggcaggtg caccagggcc tgctgaggga cgggacggat gtgggcgtg 409

<210> 294
 <211> 369
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(369)
 <223> n = A,T,C or G

<400> 294
 ggcacgaggc cagctgctgg tggagcggca ctggggactg gaggctggaa gcgggtggtg 60
 tgtgtccctt gtttactttt agctgagctg gggttgggtg tacgggttct gttcctctga 120
 gccctgcggc ccacctgatg tttacgtgtg tgtgtgaggg ggggcggngn nnnnnnnnnn 180
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnna 240
 nnnnnnnnnn ntnaatatat ttttttgttt aatgggtnnn nnnnnnnnnn nnnnnnnnnn 300
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnataaatt 360
 attaaattt 369

<210> 295
 <211> 403
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(403)
 <223> n = A,T,C or G

<400> 295
 ggcacgagtg cttctctagc tctctaggcc tctccagttt gcacctgtcc ccacctcca 60
 ctacagctgtc ctgcagcaaa cactccaccc tccaccttcc attttccccc actactgcag 120
 cacctccagg cctgttgcta tagagcctac ctgtatgtca ataaacaaca gctgaagcnn 180
 nnnnnnnnnn nnnnnccccg ccccttaaaa acaatggggg gccgtttacc gaaaacccaa 240

actggaaaaa	acccttggtg	gagttggacc	acccccacc	taaagggcgg	ggaaaaaaag	300
gctttatttg	aaaaattggg	gaggctttgg	tttaattgga	accataaaa	gccggcaaaa	360
aacaggtaac	caccaccatt	ggctttcttt	ttaggttcag	ggg		403

<210> 296

<211> 384

<212> DNA

<213> Homo sapiens

<400> 296

ggcacgagga	gaacttctgg	atcggggcca	gctcggaggc	cctcatccac	ctgggcgcca	60
agtttttcgc	ctgcatgcgc	caggaccgcg	aggtgcacag	cttcattcgc	tgggcgcgcg	120
agcgcgagaa	gcactccgcc	tgctgcgtgc	gcaacgcacg	gtcgggctgc	gtgcagacct	180
cggaggagga	gtgctcgcta	acaggaatta	tgccgtcaaa	ctcctttcca	cgttggcagt	240
gtgggtgaag	tggcccatcc	atcccagcgc	cccagagctt	gcggggccaca	agagacagtt	300
tggtctctgt	tgccaccagg	atcccagggt	gtgtgatgag	ccctcctccg	aagacctca	360
tgagtggcca	gaagacatca	ccaa				384

<210> 297

<211> 401

<212> DNA

<213> Homo sapiens

<400> 297

ggcacgagat	taagtgaatt	gcgttatatt	tatgacctta	aggaccagat	acaggaggta	60
gaaggagat	acatgcaggg	gcttaaagaa	ctaaaggaat	ctttgtctga	agtggagaa	120
aaatacaaga	aagccatggt	ttccaatgca	cagttagaca	atgagaagaa	caatttgatc	180
taccatgtag	acacactcaa	ggatgttatt	gaagagcagg	aggaacagat	ggcagaattt	240
tatagagaaa	atgaagaaaa	atcaaaggag	ttagaaaggc	agaaacatat	gtgtagtgtg	300
ctgcagcata	agatggaaga	acttaaagaa	ggcctgcggc	aaagagatga	gcttattgag	360
aaacatggct	taagtataat	ccccgatggc	actcccaatg	g		401

<210> 298

<211> 430

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(430)

<223> n = A,T,C or G

<400> 298

aaatgggaga	actctgaggt	ncccaacgat	tcgcattcgg	cacgaggcca	gctgctgggg	60
gagcggctct	ggggactgga	ggctggaagc	gggtgggtgtg	tgtcccctgt	ttacttttag	120
ctgagctggg	gttgggtgta	cgggttctgt	tcctctgagc	cctgcggccc	acctgatgtt	180
tacgtgtgtg	tgtgaggggg	ggcgggnnac	gntatanacc	catcttatta	tcaaattaca	240
aaatcccant	aataggtatc	tccatcaagc	tgcanagga	ggagagagaa	atgagagaca	300
attatgttcc	tgtgggtctca	gccttggatc	aggagattat	tgaagatgat	tcttgcctta	360
aggagatgct	gaagcttttg	gactttgggg	gtctgttcaa	ccttcattgg	acttcaactt	420
cagctgggaa						430

<210> 299

<211> 387

<212> DNA

<213> Homo sapiens

<400> 299

ggcacgaggt	ttcatacctt	ctcagaattg	gtatatcaag	acacatttaa	atataagccc	60
tctggaaatg	gatttatata	cagtcattcat	aattaccccc	ttagaaattg	gtaatatattt	120
atagccaggt	ttaggttttag	tgtcaagtat	agtgattgct	ggtctatcac	tactcatgaa	180
gtggaacccc	ctctactcat	aaaaacccc	atcagacata	tagatgaata	gaaccttgat	240
aacattagaa	tgccttggtc	tctgaaggct	tacaagacta	tacgtcagga	tatattaagg	300

agaagctgag gaacgaaaga aacttcgaca agagaatgga aatgtacatg ctatagcata 360
actgaagaat aaaatacagg tttgagg 387

<210> 300

<211> 373

<212> DNA

<213> Homo sapiens

<400> 300

ggcacgagac tagtccgact ttttatgtgc tatgcaaaat agacatcttt aacatagtcc 60
tgttactatg gtaacacttt gctttctgaa ttggaaggga aaaaaatgta acgacagcat 120
tttaagggtg ccatggtaac cagccacagt acatatgtaa ttctttccat caccccaacc 180
tctcctttct gtgcattcat gcaagagttt cttgtaagcc atcagaagtt acttttagga 240
tgggggagag gggcgagaag gggaaaaatg ggaaatagtc tgatttttaat gaaatcaaat 300
gtatgtatca tcagttggct acgttttggg tctatgctaa actgtgaaaa atcagatgaa 360
ttgataaaag agt 373

<210> 301

<211> 369

<212> DNA

<213> Homo sapiens

<400> 301

ggcacgagac tagtccgact ttttatgtgc tatgcaaaat agacatcttt aacatagtcc 60
tgttactatg gtaacacttt gctttctgaa ttggaaggga aaaaaatgta gcgacagcat 120
tttaagggtg ccatggtaac cagccacagt acatatgtaa ttctttccat caccccaacc 180
tctcctttct gtgcattcat gcaagagttt cttgtaagcc atcagaagtt acttttagga 240
tgggggagag gggcgagaag gggaaaaatg ggaaatagtc tgatttttaat gaaatcaaat 300
gtatgtatca tcagttggct acgttttggg tctatgctaa actgtgaaaa atcagatgaa 360
ttgataaaa 369

<210> 302

<211> 399

<212> DNA

<213> Homo sapiens

<400> 302

ggcacgagac agcagacacg gctgatgatg atcaaccatg accttcggca ccgggtcacg 60
gtggtggagg ccagagggaag agccctgac gaacagaagg tggagctgga ggcagacctg 120
cagaccaagg agcaggagat gggcagcctg cgagcagagc tggggaagtt gcgagagagg 180
ctgcaggggg agcacagcca gaatggggag gaggagcctg agacggagcc ggtgggagag 240
gagagcatct ccgacgcaga gaaggtggcc atggatctca aggaccccaa ccgcccccg 300
ttcaccttgc aggagctgcg ggacgtgctg cagcagagga acgagctcaa gtccaagggtg 360
ttcttgctgc aggaggagct ggcttactat aagagttag 399

<210> 303

<211> 391

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> (1)...(391)

<223> n = A,T,C or G

<400> 303

ggcacgagca cagcccctga ctgccgcagc cccacagag cccgccgcgc accccacgtc 60
ccccacgcca gcgcccagcc atggaggcca tcaagnnnnn nnnnnnnnnn nnnnnnnngg 120
acaaggagaa tgccatcgac cgcgcggagc aggcggaggc ggataagaaa gccgctgagg 180
acaagtgcaa gcaggtggag gaggagctga cgcacctcca gaagaaacta aaagggacag 240
aggacgagct ggataaatat tccgaggacc tgaaggacgc gcaggagaag ctggagctca 300
cggagaagaa ggcctccgac gctgaagggt atgtggccgc cctcaaccga cgcattccagc 360
tcgttgagga ggagttggac agggctcang a 391

<210> 304
 <211> 418
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(418)
 <223> n = A,T,C or G

<400> 304
 ggcacgagtg ccgcagcccc cacagagccc gccgcgcacc ccacgtcccc caccgccagcg 60
 cccagccatg gaggccatca agnnnnnnnn nnnnnnnnnn nnnnnggaca aggagaatgc 120
 catcgaccgc gcggagcagg cggaggcgga taagaaagcc gctgaggaca agtgcaagca 180
 ggtggaggag gagctgacgc acctccagaa gaaactaaaa gggacagagg acgagctgga 240
 taaatattcc gaggacctga aggacgcgca ggagaagctg gagctcacgg agaagaaggc 300
 ctccgacgct gaaggtgatg tggccgccct caaccgacgc atccagctcg ttgaggagga 360
 gttggacagg gctcangaac gactggccac ggccctgcag aagctggagg aggcagaa 418

<210> 305
 <211> 420
 <212> DNA
 <213> Homo sapiens

<400> 305
 ggcacgagga tttcggcaac aatttacaca gctggctgga ccagacatgg aggtgggtgc 60
 cactgatctg atgaatattc tcaacaaagt cctttctaaag cacaagatc ttaagactga 120
 cggttttagt cttgacacct gccggagcat tgtgtctgtc atggacagtg acacgactgg 180
 taagctgggc tttgaagaat ttaagtatct gtggaacaac atcaagaaat ggcagtgtgt 240
 ttataagcag tatgacaggg accattctgg gtctctggga agttctcagc tgcggggagc 300
 tctgcaggcc gcaggcttcc agctaaatga acaactttac caaatgattg tccgccggta 360
 tgctaataaa gatggagata tggattttaa caatttcac agctgcttgg tccgcctgga 420

<210> 306
 <211> 399
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(399)
 <223> n = A,T,C or G

<400> 306
 ggcacgagcc acgtccccca cgccagcgcc cagccatgga ggccatcaag nnnnnnnnnn 60
 nnnnnnnnnn nnnnggacaag gagaatgcca tcgaccgcgc ggagcaggcg gaggcggata 120
 agaaagccgc tgaggacaag tgcaagcagg tggaggagga gctgacgcac ctccagaaga 180
 aactaaaagg gacagaggac gagctggata aatattccga ggacctgaag gacgcgcagg 240
 agaagctgga gctcacggag aagaaggcct ccgacgctga aggtgatgtg gccgccctca 300
 accgacgcat ccagctcgtt gaggaggagt tggacagggc tcaggaacga ctggccacgg 360
 ccctgcagaa gctggaggag gcagaanaag ctgcagatg 399

<210> 307
 <211> 438
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(438)
 <223> n = A,T,C or G

<400> 307

atcccatcga	ttcgaattcg	gcacgagccc	ccacagagcc	cgccgtgcac	cccacgtccc	60
ccacgccagc	gcccagccat	ggaggccatc	aagnnnnnnn	nnnnnnnnnn	nnnnnnnggac	120
aaggagaatg	ccatcgaccg	cgcggagcag	gcggaggcgg	ataagaaagc	cgctgaggac	180
aagtgcgaagc	aggtggagga	ggagctgacg	cacctccaga	agaaactaaa	agggacagag	240
gacgagctgg	ataaatattc	cgaggacctg	aaggacgcgc	aggagaagct	ggagctcacg	300
gagaagaagg	cctccgacgc	tgaaggatgat	gtggccgccc	tcaaccgacg	catccagctc	360
gttgaggagg	agttggacag	ggctcatgaa	cgactgggca	cggacctgca	gaagctggag	420
gagggcagaa	aaagctgc					438

<210> 308

<211> 419

<212> DNA

<213> Homo sapiens

<400> 308

ggcagcagct	ttggcctgcc	cgctcctctc	ctttctggcg	acccgactct	ggctacgcaa	60
cggggcccgc	gtcaatgcct	gggcctactg	ccacgtgcta	cccactgggg	acctgctgct	120
ggtgggcacc	caacagctgg	gggagttcca	gtgctggtca	ctagaggagg	gcttccagca	180
gctggtagcc	agctactgcc	cacaggtggg	ggaggacggc	gtggcagacc	aaacagatga	240
gggtggcagt	gtacccgtca	ttatcagcac	atcgcggtgtg	agtgcaccac	ctgggtggcaa	300
ggccagctgg	ggtgcagaca	ggtcctactg	gaaggagttc	ctggtgatgt	gcacgctctt	360
tgtgctggcc	gtgctgctcc	cagttttatt	cttgctctac	cggcacccgga	acagcatgg	419

<210> 309

<211> 415

<212> DNA

<213> Homo sapiens

<400> 309

ggcagcaggg	tgagccagag	acgccctcca	ttctctcttc	gcgcccgcgc	tccggtgggc	60
ctcccgatgc	gctgcccgcg	ctgccaccat	gacggaacag	gccatctcct	tcgacaaaga	120
cttcttggcc	ggaggcatcg	tcgccgtcat	cttcaagacg	gacgtggctc	ctatcgagcg	180
ggtcaagctg	ctgctgccgt	ccagcacgcc	agcaagcaga	tcgcccgcga	ctagcagtac	240
aagggcatcg	tggactgcat	tgtccgcac	cccaaagagc	atggagtgtc	gtccttctgg	300
aagggcaacc	ttgccaacgt	caatcgctac	ttccccactc	aagccctcaa	cttcgtcttc	360
aaggataatg	acatgcagat	cttactgggg	ggcgtggaca	aacacacgca	ggtct	415

<210> 310

<211> 396

<212> DNA

<213> Homo sapiens

<400> 310

ggcagcagcg	ggtcctgccg	gtgccacatg	gggtaccagg	gcccgcgtgtg	cactgactgc	60
atggacggct	acttcagctc	gctccggaac	gagaccacac	gcatctgcac	agcctgtgac	120
gagtcctgca	agacgtgtc	gggcctgacc	aacagagact	gcggcgagtgtg	tgaagtgggc	180
tgggtgtctg	acgagggcgc	ctgtgtggat	gtggacgagt	gtgcggccga	gccgcctccc	240
tgcagcgctg	cgcagttctg	taagaacgcc	aacggctcct	acacgtgcga	agagtgtgac	300
tccagctgtg	tgggctgcac	aggggaaggc	ccaggaaact	gtaaagagtgtg	tatctctggc	360
tacgcgaggg	agcacggaca	gtgtgcagat	gtggac			396

<210> 311

<211> 394

<212> DNA

<213> Homo sapiens

<400> 311

ggcagcaggg	ctctggggccc	tacagctcat	cctggtcacg	tgccctcac	tgtcgtgggt	60
catgcacgtg	gcctaccgcg	aggaacgcga	gcgcaagcac	cacctgaaac	acgggcccaa	120
tgccccgtcc	ctgtacgaca	acctgagcaa	gaagcggggc	ggactgtgggt	ggacgtactt	180
gctgagcctc	atcttcaagg	ccgccgtgga	tgtgtgcttc	ctctatatct	tccaccgcct	240
ctacaaggat	tatgacatgc	cccgcgtggg	ggcctgtctc	gtggagcctt	gccccacac	300
tgtggactgt	tacatctccc	ggcccacgga	gaagaaggtc	ttcacctact	tcatggtgac	360

cacagctgca tggagatctt cggccccagg cacc 394

<210> 312
 <211> 384
 <212> DNA
 <213> Homo sapiens

<400> 312
 ggcacgaggc gaggaacgcg agcgcaagca ccacctgaaa caggggcca atgccccgtc 60
 cctgtacgac aacctgagca agaagcggg cggactgtgg tggacgtact tgctgagcct 120
 catcttcaag gccgcccgtg atgctggctt cctctatata tccaccgcc tctacaagga 180
 ttatgacatg ccccgctgg tggcctgctc cgtggagcct tgccccaca ctgtggactg 240
 ttacatctcc cggcccacgg agaagaaggt cttcacctac ttcattgtga ccacagctgc 300
 catctgcata ctgctcaacc tcagtgaagt cttctacctg gtgggcaaga ggtgcatgga 360
 gatcttcggc cccaggcacc ggcg 384

<210> 313
 <211> 430
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(430)
 <223> n = A,T,C or G

<400> 313
 ggcacgagcc ggctcgtaag caacctcttc agtctgcagt gggacccgcg cgtcatgcag 60
 cgtgccagca gcaacctgca ccgcggtccg ggcggggcgc tggctcttct ggacaatgag 120
 gcgggcttgg tgcacggcta ccgggtagca ggcattgtgg acaagtataa cgagccgctg 180
 ttgcagtcat tgtgcgtgtt ccgcgagcgg accgcgcggc gcgtcctgga gctgcaccgc 240
 ggacaggacg ccgcggcccg gctgctgctc ctctaccggc gccacgagcc tcgcttcccc 300
 gagctggccg cccttgacga ccccccacgt cagctgctac agcgccgct cgaattcctc 360
 gccaaacaca ttttgactg taaggccaag tacggccgcc ggtctgggac ttagtgtcac 420
 cgggaggaan 430

<210> 314
 <211> 408
 <212> DNA
 <213> Homo sapiens

<400> 314
 ggcacgagag cagaaggact ttgtctgcaa caccaagcag cccggctgcc ccaacgtctg 60
 ctatgacgag ttcttccccg tgtcccacgt gcgcctctgg gccctacagc tcactcctggt 120
 cagtgcccc tcactgctcg tggctatgca cgtggcctac cgcgaggaac gcgagcgcaa 180
 gcaccacctg aaacacgggc ccaatgcccc gtccctgtac gacaacctga gcaagaagcg 240
 gggcggactg tgggtggacgt acttgctgag cctcatcttc aaggccgccc tggatgctgg 300
 ctctctctat atcttccacc gcctctacaa ggattatgac atgccccgcg tgggtggcctg 360
 ctccgtggag ccttgccccc acactgtgga ctgttacatc tcccggcc 408

<210> 315
 <211> 412
 <212> DNA
 <213> homo sapiens

<400> 315
 tcggagccca tgcgcagcgg ggcgcgttag ctgcgcgtct tcctgacccc cgatcctggg 60
 gccgaggtac ctttgacagg agcgtgaccc tgctggaggt gtgcgggagc tggcctgagg 120
 gcttcgggct gcggcacatg tcctccatgg agcacacgga ggaggccctc cgggagcgac 180
 ttgccgacgc catggccgag tcacctagcc gggacgtcgt gggatccgga acagaacttc 240
 agcgagaggg aagcatcgag actctgagta acagctcagg ctccaccagc ggcagcatac 300
 caagaaactt tgatggctac cgatctccgc tgcccaccaa tgagagccag cccctcagcc 360
 tcttcccagc tggcttcccc taggtaccag caacctgctt ctgactggcc ag 412

<210> 316
 <211> 300
 <212> DNA
 <213> homo sapiens

<400> 316
 gccagccctt cagcctcttc ccgactggct tcccgtaggt accagcaacc tgctttctgac 60
 tggccagccc cctcccctgc tggaggagg gagaagcccc gctctgggtc tacccttcag 120
 tctctgctct tccttcatca accaccttcc ccaagcttag tgacagcagc cgcccatcct 180
 acctggatgg agaagagacc cttctccaag cacctcagcg cacttgcctt ctgccacacc 240
 tgtcgggtgga ggctgtggcc aggagagact gtagaagctc ggtccctgtg tatgttttga 300

<210> 317
 <211> 2064
 <212> DNA
 <213> homo sapiens

<400> 317
 acctcagcca gattcggcac gaggggcgta ggaccctccg agccagggtgt gggatatagt 60
 ctctgtggtgc gccgtttttt aagccggtct gaaaagcgca atattcgggt gggagtgacc 120
 cgattttcca ggctgctatc catgtccagg gccaaacatg aatcctattg ctcttgaggga 180
 gccgctgggt tgcttatgca gaaaacaagt tgattcgatg tcatcagtcg cgtgggtggag 240
 cctgtggaga caacattcag tcttatactg ccacagtcac tagtgctgct aaaacattga 300
 aaagtggcct gacaatggta gggaaagtgg tgactcagct gacaggcaca ctgccttcag 360
 gtgtgacaga agatgatgtt gccatccaca gtaattcacg gcggagtcct ttgggtcccag 420
 gcatcatcac agttattgac accgaaaccg ttggagaggg ccagggtgctt gtgagtgagg 480
 attctgacag tgatggcatt gtggccactt tccctgcccc tgagaagcca gtgtgctgca 540
 tggcttttaa tacaagtgga atgcttctag tcacaacaga cacccttggc catgactttc 600
 atgtcttcca aattctgact catccttggg cctcatcaca atgtgctgtc caccatctgt 660
 atactcttca caggggagaa actgaagcca aagtacagga catctgcttc agccatgact 720
 gtcgctgggt tgtggtcagt actctccggg gtacttccca cgttttcccc atcaaccctt 780
 atggtggcca gccttgtgtt cgtacacata tgtcaccacg agtagtgaat cgcattgagcc 840
 gtttccagaa aagtgtctga ctggaagaga ttgaacaaga actgacgtct aagcaaggag 900
 gtcgctgtag ccctgttcca ggtctatcaa gcagcccttc tgggtcaccg ttgcatggga 960
 aactgaacag ccaagactcc tataacaatt ttaccaacaa caaccctggc aaccctcggc 1020
 tctctcctct tcccagcttg atggtagtga tgcctcttgc acaaatcaag cagccaatga 1080
 cattggggag catcaccaa cgaaccggca aagttaaacc tcctccacaa atttcaccca 1140
 gcaaatcgat gggcggagaa ttttgtgtgg ctgctatctt cggaacatcc aggtcatggg 1200
 ttgcaataaa tgcaggctctg aaaagagaaa aagatcagtc caaacaagtt gtagttgagt 1260
 ccctgtacat tatcagttgc tatggcacct tagtggaaac catgatggag ccgcgacccc 1320
 tcagcactgc acccaagatt agtgacgaca caccactgga aatgatgaca tgcctcgag 1380
 ccagctggac tctgggttaga acccctcaat ggaatgaatt gcagccaccg tttaatgcaa 1440
 accaccctct gctcctcgct gcagatgcag tacagtatta tcagttcctg ctgtctggcc 1500
 tggttccccc ttggaagtcct gggcccattg ctgcacatgg gtcttacgac gttctagctt 1560
 ctgaccatag tggacaggaa gatgaagaat ggctttccca ggttgaaatt gtaacacaca 1620
 ctggacccca tagacgtctg tggatgggtc cacagttcca gttcaaaacc atccatccct 1680
 caggccaaac cacagttatc tcatccagtt catctgtgtt gcagtctcat ggtccgagtg 1740
 acacgccaca gcctcttttg gattttgata cagatgatct tgatctcaac agtctcagga 1800
 tccagccagt ccgctctgac cccgtcagca tgccagggtc atcccgtcca gtctctgac 1860
 gaaggggagt ttccacagtg attgatgctg cctcaggtag ctttgacagg agcgtgaccc 1920
 tgctggagggt gtgcgggagc tggcctgagg gcttcgggct cgggcacatg tctccatgg 1980
 agcacacgga ggagggcctc cgggagcgac ttgccgacgc catggccgag tcacctagcc 2040
 gggacgtcgt gggatccgga acag 2064

<210> 318
 <211> 1365
 <212> DNA
 <213> homo sapiens

<400> 318
 cgagaactct gagtaacagc tcagggtcca ccagcggcag cataccaaga aactttgatg 60
 gctaccgate tccgctgccc accaatgaga gccagccctt cagcctcttc ccgactggct 120
 tcccgtaggt accagcaacc tgcttctgac tggccagccc cctcccctgc tggaggaggg 180

gagaagcccc	gctctggtcc	tacccttcag	tctctgctct	tccttcatca	accaccttcc	240
ccaagcttag	tgacagcagc	cgcccatcct	acctggatgg	agaagagacc	cttctccaag	300
cacctcagcg	cacttgccct	ctgccacacc	tgctcgggtga	ggctgtggcc	aggagagact	360
gtagaagctc	ggtccctgtg	tatgtttgca	tatgacatcc	tgcataggat	ccgcttttgt	420
atTTTTTaa	cataccacg	gtggggcg	tggggggagc	ctggaacagt	gaccagatct	480
gggggcctga	gtgggggacag	agttgatcgt	ccacctggcc	atTTTgacc	tgagtggaca	540
gtcacagcct	cagctcatgt	ctggctgtga	cacacactgc	ccccagcttc	ccttggtcag	600
ccccactcca	gcacggggtg	aacggaggcc	cagagtacta	gggaaggagg	aaggaggagc	660
atgcctcttc	ttcctccttt	ctttcccat	ctgttcctgg	gaagagtttg	tctttcttat	720
ctttaagccc	ctttaccctg	gtcctgtact	gatcagtga	ggaaaccgtg	gttactgagg	780
ccctgttgaa	aagtgcacgt	cttgtccaat	aaatcacgct	gcagttggaa	aaaaaaaaa	840
aaaaaaaaa	gatctttaat	taagcgccg	caagcttatt	ccctttagtg	aggggtaatt	900
ttagcttggc	actggccgtc	gttttacaac	gtcgtgactg	gtaaaccctg	gcgttaccga	960
acttaatcgc	cttgacgac	atcccccttt	cgccagctgg	cgtaatagcg	aagaggcccg	1020
caccgatcgc	ccttcccaac	agttgcgcag	cctgaatggc	gaatgggacg	cgccctgtag	1080
cggcgcatga	agcgcggcgg	gtgtggtggt	taccgcgcag	cgtgaccgct	acacttgcca	1140
gcgccttagc	gcccgcctct	ttcgctttct	tcccccttct	ttttcgccac	gttcgcccgc	1200
tttccccctg	caagctctaa	atcgggggct	cccccttagg	gttcccgatt	tagtgcttta	1260
cgggcacctc	gacccccaaa	aacttgatta	gggtgatggt	tcacgtagtg	ggccatcgcc	1320
ctgataagac	ggTTTTTcgc	cctttgacgt	tgaggtccac	gttct		1365

<210> 319

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthesized primer

<400> 319

tgggatatag tctcgtggtg cg

22

<210> 320

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthesized primer

<400> 320

tgattcgatg tcatcagtcg cg

22

<210> 321

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> synthesized primer

<400> 321

tgtgtcacag ccagacatga gc

22

<210> 322

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> synthesized primer

<400> 322

tgcaaacata cacagggacc g 21

<210> 323
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> synthesized primer

<400> 323
tttagcagca ctaatgactg tggc 24

<210> 324
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> synthesized primer

<400> 324
cgccgtgaat tactgtggat gg 22